

# SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

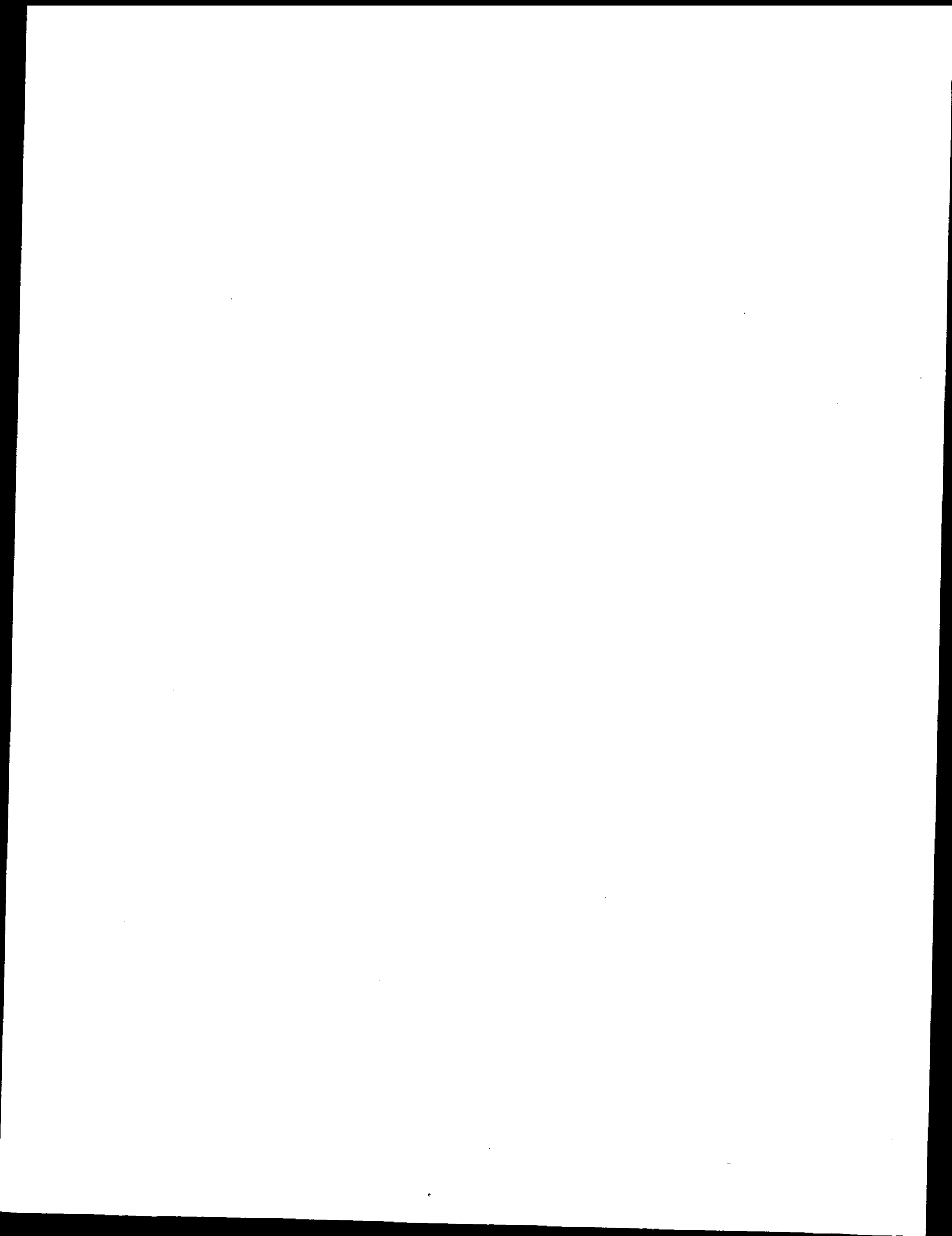
## STAFF USE ONLY

Date completed: 08-28-03  
Searcher: Beverly C 4994  
Terminal time: 20  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 25  
Number of Searches: \_\_\_\_\_  
Number of Databases: 1

Search Site  
\_\_\_\_ STIC  
\_\_\_\_ CM-1  
\_\_\_\_ Pre-S  
Type of Search  
\_\_\_\_ N.A. Sequence  
\_\_\_\_ A.A. Sequence  
\_\_\_\_ Structure  
\_\_\_\_ Bibliographic

Vendors  
\_\_\_\_ IG  
\_\_\_\_ STN  
\_\_\_\_ Dialog  
\_\_\_\_ APS  
\_\_\_\_ Geninfo  
\_\_\_\_ SDC  
\_\_\_\_ DARC/Questel  
\_\_\_\_ Other CGN

BEST AVAILABLE COPY



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 27, 2003, 00:51:44; Search time 130.55 Seconds  
(without alignments)  
7529.936 Million cell updates/sec

Title: US-09-937-982-1  
Perfect score: 30  
Sequence: 1 tccggatccagcgcctctgtttgatggt 30  
Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0  
Searched: 32997241 seqs, 16383922548 residues  
Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Pending patents NA Main:
- 1: /cgn2\_6/ptodata/1/pna/PCTUS\_COMB.seq.\*
  - 2: /cgn2\_6/ptodata/1/pna/PCTUS\_COMB.seq.\*
  - 3: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq.\*
  - 4: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq.\*
  - 5: /cgn2\_6/ptodata/1/pna/US08\_COMB.seq.\*
  - 6: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq.\*
  - 7: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq.\*
  - 8: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq.\*
  - 9: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq.\*
  - 10: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq.\*
  - 11: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq.\*
  - 12: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq.\*
  - 13: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq.\*
  - 14: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq.\*
  - 15: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq.\*
  - 16: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq.\*
  - 17: /cgn2\_6/ptodata/1/pna/US092A\_COMB.seq.\*
  - 18: /cgn2\_6/ptodata/1/pna/US092B\_COMB.seq.\*
  - 19: /cgn2\_6/ptodata/1/pna/US093A\_COMB.seq.\*
  - 20: /cgn2\_6/ptodata/1/pna/US093B\_COMB.seq.\*
  - 21: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq.\*
  - 22: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq.\*
  - 23: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq.\*
  - 24: /cgn2\_6/ptodata/1/pna/US095C\_COMB.seq.\*
  - 25: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq.\*
  - 26: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq.\*
  - 27: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq.\*
  - 28: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq.\*
  - 29: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq.\*
  - 30: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq.\*
  - 31: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq.\*
  - 32: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq.\*
  - 33: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq.\*
  - 34: /cgn2\_6/ptodata/1/pna/US098A\_COMB.seq.\*
  - 35: /cgn2\_6/ptodata/1/pna/US098B\_COMB.seq.\*
  - 36: /cgn2\_6/ptodata/1/pna/US098C\_COMB.seq.\*
  - 37: /cgn2\_6/ptodata/1/pna/US098D\_COMB.seq.\*
  - 38: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq.\*
  - 39: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq.\*
  - 40: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq.\*
  - 41: /cgn2\_6/ptodata/1/pna/US099D\_COMB.seq.\*
  - 42: /cgn2\_6/ptodata/1/pna/US099E\_COMB.seq.\*
  - 43: /cgn2\_6/ptodata/1/pna/US099F\_COMB.seq.\*

- 44: /cgn2\_6/ptodata/1/pna/US100A\_COMB.seq.\*
- 45: /cgn2\_6/ptodata/1/pna/US100B\_COMB.seq.\*
- 46: /cgn2\_6/ptodata/1/pna/US101A\_COMB.seq.\*
- 47: /cgn2\_6/ptodata/1/pna/US101B\_COMB.seq.\*
- 48: /cgn2\_6/ptodata/1/pna/US102A\_COMB.seq.\*
- 49: /cgn2\_6/ptodata/1/pna/US102B\_COMB.seq.\*
- 50: /cgn2\_6/ptodata/1/pna/US103A\_COMB.seq.\*
- 51: /cgn2\_6/ptodata/1/pna/US103B\_COMB.seq.\*
- 52: /cgn2\_6/ptodata/1/pna/US104A\_COMB.seq.\*
- 53: /cgn2\_6/ptodata/1/pna/US104B\_COMB.seq.\*
- 54: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq.\*
- 55: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq.\*
- 56: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq.\*
- 57: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq.\*
- 58: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq.\*
- 59: /cgn2\_6/ptodata/1/pna/US6005\_COMB.seq.\*
- 60: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq.\*
- 61: /cgn2\_6/ptodata/1/pna/US6007\_COMB.seq.\*
- 62: /cgn2\_6/ptodata/1/pna/US6008\_COMB.seq.\*
- 63: /cgn2\_6/ptodata/1/pna/US6009\_COMB.seq.\*
- 64: /cgn2\_6/ptodata/1/pna/US6010\_COMB.seq.\*
- 65: /cgn2\_6/ptodata/1/pna/US6011\_COMB.seq.\*
- 66: /cgn2\_6/ptodata/1/pna/US6012\_COMB.seq.\*
- 67: /cgn2\_6/ptodata/1/pna/US6013\_COMB.seq.\*
- 68: /cgn2\_6/ptodata/1/pna/US6014\_COMB.seq.\*
- 69: /cgn2\_6/ptodata/1/pna/US6015\_COMB.seq.\*
- 70: /cgn2\_6/ptodata/1/pna/US6016\_COMB.seq.\*
- 71: /cgn2\_6/ptodata/1/pna/US6017\_COMB.seq.\*
- 72: /cgn2\_6/ptodata/1/pna/US6018\_COMB.seq.\*
- 73: /cgn2\_6/ptodata/1/pna/US6019\_COMB.seq.\*
- 74: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq.\*
- 75: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq.\*
- 76: /cgn2\_6/ptodata/1/pna/US6022\_COMB.seq.\*
- 77: /cgn2\_6/ptodata/1/pna/US6023A\_COMB.seq.\*
- 78: /cgn2\_6/ptodata/1/pna/US6023B\_COMB.seq.\*
- 79: /cgn2\_6/ptodata/1/pna/US6024\_COMB.seq.\*
- 80: /cgn2\_6/ptodata/1/pna/US6025\_COMB.seq.\*
- 81: /cgn2\_6/ptodata/1/pna/US6026\_COMB.seq.\*
- 82: /cgn2\_6/ptodata/1/pna/US6027\_COMB.seq.\*
- 83: /cgn2\_6/ptodata/1/pna/US6028\_COMB.seq.\*
- 84: /cgn2\_6/ptodata/1/pna/US6029\_COMB.seq.\*
- 85: /cgn2\_6/ptodata/1/pna/US6030\_COMB.seq.\*
- 86: /cgn2\_6/ptodata/1/pna/US6031\_COMB.seq.\*
- 87: /cgn2\_6/ptodata/1/pna/US6032\_COMB.seq.\*
- 88: /cgn2\_6/ptodata/1/pna/US6033\_COMB.seq.\*
- 89: /cgn2\_6/ptodata/1/pna/US6034\_COMB.seq.\*
- 90: /cgn2\_6/ptodata/1/pna/US6035\_COMB.seq.\*
- 91: /cgn2\_6/ptodata/1/pna/US6036\_COMB.seq.\*
- 92: /cgn2\_6/ptodata/1/pna/US6037\_COMB.seq.\*
- 93: /cgn2\_6/ptodata/1/pna/US6038\_COMB.seq.\*
- 94: /cgn2\_6/ptodata/1/pna/US6039\_COMB.seq.\*
- 95: /cgn2\_6/ptodata/1/pna/US6040\_COMB.seq.\*
- 96: /cgn2\_6/ptodata/1/pna/US6041\_COMB.seq.\*
- 97: /cgn2\_6/ptodata/1/pna/US6042\_COMB.seq.\*
- 98: /cgn2\_6/ptodata/1/pna/US6043\_COMB.seq.\*
- 99: /cgn2\_6/ptodata/1/pna/US6044\_COMB.seq.\*
- 100: /cgn2\_6/ptodata/1/pna/US6045\_COMB.seq.\*
- 101: /cgn2\_6/ptodata/1/pna/US6046\_COMB.seq.\*
- 102: /cgn2\_6/ptodata/1/pna/US6047\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	US-09-773-260-1	Sequence 1, Appli
2	30	100.0	30	US-09-937-982-1	Sequence 1, Appli
3	28	93.3	1133	US-09-773-260-3	Sequence 3, Appli
4	28	93.3	1133	US-09-937-982-3	Sequence 3, Appli

5 21.6 72.0 1836 53 US-10-437-963-45955  
6 21.6 72.0 35740 27 US-09-620-392-50458  
7 21.6 72.0 35740 31 US-09-702-134-25459  
8 21.6 72.0 35740 34 US-09-815-264-71453  
9 20.8 69.3 1694969 91 US-60-361-742-1692  
10 20.6 68.7 1287 21 US-09-417-507-19466  
11 20.4 68.0 300 37 US-09-881-797-4336  
12 20.4 68.0 300 44 US-10-021-698-4336  
13 20.4 68.0 599 22 US-09-505-532-29293  
14 20.4 68.0 599 34 US-09-534-859-763  
15 20.4 68.0 5015 23 US-09-803-736-763  
16 20.4 68.0 80741 101 US-60-465-241-52917  
17 20.4 68.0 80741 101 US-60-466-412-86873  
18 20.4 66.7 502 18 US-09-270-849B-59175  
19 20.4 66.7 599 28 US-09-654-617-184307  
20 20.4 66.7 599 30 US-09-684-016-184307  
21 19.8 66.0 573 20 US-09-371-168-4975  
22 19.8 66.0 573 34 US-09-801-833-4975  
23 19.8 66.0 697 20 US-09-371-168-7927  
24 19.8 66.0 697 34 US-09-801-833-7927  
25 19.8 66.0 3534 80 US-60-258-273-214  
26 19.8 66.0 6802 40 US-09-949-016-2442  
27 19.8 66.0 6868 38 US-09-919-039-39  
28 19.8 66.0 6868 46 US-10-133-013-139  
29 19.8 66.0 6902 28 US-09-654-788A-130  
30 19.8 66.0 6902 49 US-10-272-268-130  
31 19.8 66.0 6906 1 PCT-US02-21338-77  
32 19.8 66.0 6906 1 PCT-US02-21338-77  
33 19.8 66.0 6906 47 US-10-188-832-77  
34 19.8 66.0 6981 1 PCT-US01-24708-636  
35 19.8 66.0 6981 1 PCT-US02-27826-299  
36 19.8 66.0 6981 1 PCT-US02-10873A-255  
37 19.8 66.0 6981 1 PCT-US02-25766-2609  
38 19.8 66.0 6981 2 PCT-US02-30182-977  
39 19.8 66.0 6981 2 PCT-US01-24708-636  
40 19.8 66.0 6981 36 US-09-873-319-636  
41 19.8 66.0 6981 42 US-09-960-706-977  
42 19.8 66.0 6981 42 US-09-968-007A-424

ALIGNMENTS

RESULT 1  
US-09-773-260-1  
; Sequence 1, Application US/09773260  
; GENERAL INFORMATION:  
; APPLICANT: Donald L Durden  
; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES  
; ASPARAGINASE IN THE TREATMENT OF HUMAN  
; HEMATOLOGIC AND AUTOIMMUNE DISEASE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; CITY: Suite 4700  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq for Windows 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/773,260  
; FILING DATE: 31-Jan-2001  
; CLASSIFICATION: <Unknown>

Sequence 45955, A  
Sequence 50458, A  
Sequence 25459, A  
Sequence 71453, A  
Sequence 1692, Ap  
Sequence 19466, A  
Sequence 4336, Ap  
Sequence 4336, Ap  
Sequence 29293, A  
Sequence 29293, A  
Sequence 763, App  
Sequence 763, App  
Sequence 52917, A  
Sequence 86873, A  
Sequence 59175, A  
Sequence 184307, A  
Sequence 4975, Ap  
Sequence 4975, Ap  
Sequence 7927, Ap  
Sequence 7927, Ap  
Sequence 214, App  
Sequence 2442, App  
Sequence 39, Appl  
Sequence 139, App  
Sequence 39, Appl  
Sequence 130, App  
Sequence 130, App  
Sequence 130, App  
Sequence 77, Appl  
Sequence 77, Appl  
Sequence 636, App  
Sequence 299, App  
Sequence 255, App  
Sequence 2609, App  
Sequence 977, App  
Sequence 636, App  
Sequence 977, App  
Sequence 424, App

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/094,435  
FILING DATE: 1998-06-09  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 234/274  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-773-260-1

Query Match 100.0%; Score 30; DB 33; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCCGGATCCAGCGCCCTCTGTTTGGATGGCT 30  
Db 1 TCCGGATCCAGCGCCCTCTGTTTGGATGGCT 30

RESULT 2  
US-09-937-982-1  
; Sequence 1, Application US/09937982  
; GENERAL INFORMATION:  
; APPLICANT: DURDEN, Donald L  
; TITLE OF INVENTION: DISEASE AND GRAFT VERSUS HOST DISEASE  
; FILE REFERENCE: 270/29905  
; CURRENT APPLICATION NUMBER: US/09/937,982  
; CURRENT FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: PCT/US00/07981  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: US 60/127,662  
; PRIOR FILING DATE: 1999-04-02  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 30  
; TYPE: DNA  
; ORGANISM: Wolinella succinogenes  
US-09-937-982-1

Query Match 100.0%; Score 30; DB 39; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TCCGGATCCAGCGCCCTCTGTTTGGATGGCT 30  
Db 1 TCCGGATCCAGCGCCCTCTGTTTGGATGGCT 30

RESULT 3  
US-09-773-260-3  
; Sequence 3, Application US/09773260  
; GENERAL INFORMATION:  
; APPLICANT: Donald L Durden  
; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES  
; ASPARAGINASE IN THE TREATMENT OF HUMAN  
; HEMATOLOGIC AND AUTOIMMUNE DISEASE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street

; Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071-2066  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ;  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: FastSeq for Windows 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/773,260  
 ; FILING DATE: 31-Jan-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/094,435  
 ; FILING DATE: 1998-06-09  
 ; APPLICATION NUMBER: <Unknown>  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Warburg, Richard J.  
 ; REGISTRATION NUMBER: 32,327  
 ; REFERENCE/DOCKET NUMBER: 234/274  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ;  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1133 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
 US-09-773-260-3

Query Match 93.3%; Score 28; DB 39; Length 1133;  
 Best Local Similarity 100.0%; Pred. No. 0.18;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTGTTTGTATGGCT 30  
 DB 96 CGGATCCAGCGCCTCTGTTTGTATGGCT 123

RESULT 4  
 ; US-09-937-982-3  
 ; Sequence 3, Application US/09937982  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DURDEN, Donald L  
 ; TITLE OF INVENTION: DISEASE AND GRAFT VERSUS HOST DISEASE  
 ; FILE REFERENCE: 270/299US  
 ; CURRENT APPLICATION NUMBER: US/09/937,982  
 ; CURRENT FILING DATE: 2002-03-22  
 ; PRIOR APPLICATION NUMBER: PCT/US00/07981  
 ; PRIOR FILING DATE: 2000-03-24  
 ; PRIOR APPLICATION NUMBER: US 60/127,662  
 ; PRIOR FILING DATE: 1999-04-02  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 3  
 ; LENGTH: 1133  
 ; TYPE: DNA  
 ; ORGANISM: Wolinella succinogenes  
 US-09-937-982-3

Query Match 93.3%; Score 28; DB 39; Length 1133;  
 Best Local Similarity 100.0%; Pred. No. 0.18;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTGTTTGTATGGCT 30  
 DB 96 CGGATCCAGCGCCTCTGTTTGTATGGCT 123

Db 96 CGGATCCAGCGCCTCTGTTTGTATGGCT 123

# RESULT 5

; US-10-437-963-45955  
 ; Sequence 45955, Application US/10437963  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 45955  
 ; LENGTH: 1836  
 ; TYPE: DNA  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_4886C.1  
 US-10-437-963-45955

Query Match 72.0%; Score 21.6; DB 53; Length 1836;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTGTTTGTATGGCT 30  
 DB 1672 CGCCTCTGCTCTCTGTTTGTATGGCT 1699

# RESULT 6

; US-09-620-392-50458  
 ; Sequence 50458, Application US/09620392  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: McIninch, James  
 ; TITLE OF INVENTION: Plant Genome Sequence And Uses Thereof  
 ; FILE REFERENCE: 38-21(51237)E  
 ; CURRENT APPLICATION NUMBER: US/09/620,392  
 ; CURRENT FILING DATE: 2000-07-19  
 ; NUMBER OF SEQ ID NOS: 69652  
 ; SEQ ID NO 50458  
 ; LENGTH: 35740  
 ; TYPE: DNA  
 ; ORGANISM: Oryza sativa  
 ; OTHER INFORMATION: unsure at all n locations  
 US-09-620-392-50458

Query Match 72.0%; Score 21.6; DB 27; Length 35740;  
 Best Local Similarity 85.7%; Pred. No. 1.7e+02;  
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTGTTTGTATGGCT 30  
 DB 13407 CGCCTCTGCTCTCTGTTTGTATGGCT 13434

# RESULT 7

; US-09-702-134-25459  
 ; Sequence 25459, Application US/09702134  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Cao, Yongwei

ALLIANCE: fidelity systems, inc.

us-09-937-982-1.1.rnmpm

Thu Aug 28 09:00:59 2003

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-09-881-797-4336

Query Match      68.0%; Score 20.4; DB 37; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTTGATGG 167

RESULT 12
US-10-021-698-4336/c
; Sequence 4336, Application US/10021698
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; FILE REFERENCE: HUM01-04
; CURRENT APPLICATION NUMBER: US/10/021,698
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4687
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4336
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-10-021-698-4336

Query Match      68.0%; Score 20.4; DB 44; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTTGATGG 167

RESULT 13
US-09-505-532-29293/c
; Sequence 29293, Application US/09505532
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Plant Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15478)B
; CURRENT APPLICATION NUMBER: US/09/505,532
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: US/09/505,532
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 51470
; SEQ ID NO 29293
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
US-09-505-532-29293

Query Match      68.0%; Score 20.4; DB 22; Length 599;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTCTTTTGATGGCT 30
    |||||
Db 188 TCCGGTTTCATCGCCACTGTTTGTGATCACT 159

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-09-881-797-4336

Query Match      68.0%; Score 20.4; DB 37; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTTGATGG 167

RESULT 12
US-10-021-698-4336/c
; Sequence 4336, Application US/10021698
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; FILE REFERENCE: HUM01-04
; CURRENT APPLICATION NUMBER: US/10/021,698
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4687
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4336
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-10-021-698-4336

Query Match      68.0%; Score 20.4; DB 44; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTTGATGG 167

RESULT 13
US-09-505-532-29293/c
; Sequence 29293, Application US/09505532
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Plant Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15478)B
; CURRENT APPLICATION NUMBER: US/09/505,532
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: US/09/505,532
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 51470
; SEQ ID NO 29293
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
US-09-505-532-29293

Query Match      68.0%; Score 20.4; DB 22; Length 599;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTCTTTTGATGGCT 30
    |||||
Db 188 TCCGGTTTCATCGCCACTGTTTGTGATCACT 159

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-09-881-797-4336

Query Match      68.0%; Score 20.4; DB 37; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTTGATGG 167

RESULT 12
US-10-021-698-4336/c
; Sequence 4336, Application US/10021698
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; FILE REFERENCE: HUM01-04
; CURRENT APPLICATION NUMBER: US/10/021,698
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4687
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4336
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-10-021-698-4336

Query Match      68.0%; Score 20.4; DB 44; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTTGATGG 167

RESULT 13
US-09-505-532-29293/c
; Sequence 29293, Application US/09505532
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Plant Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15478)B
; CURRENT APPLICATION NUMBER: US/09/505,532
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: US/09/505,532
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 51470
; SEQ ID NO 29293
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
US-09-505-532-29293

Query Match      68.0%; Score 20.4; DB 22; Length 599;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTCTTTTGATGGCT 30
    |||||
Db 188 TCCGGTTTCATCGCCACTGTTTGTGATCACT 159

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-09-881-797-4336

Query Match      68.0%; Score 20.4; DB 37; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTTGATGG 167

RESULT 12
US-10-021-698-4336/c
; Sequence 4336, Application US/10021698
; GENERAL INFORMATION:
; APPLICANT: Tim Keith
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; FILE REFERENCE: HUM01-04
; CURRENT APPLICATION NUMBER: US/10/021,698
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 4687
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4336
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(300)
; OTHER INFORMATION: n = A,T,C or G
US-10-021-698-4336

Query Match      68.0%; Score 20.4; DB 44; Length 300;
Best Local Similarity 91.3%; Pred. No. 3.4e+02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTCTTTTGATGG 28
    |||||
Db 189 ATCCAGCGCCTCTCTTTTGATGG 167

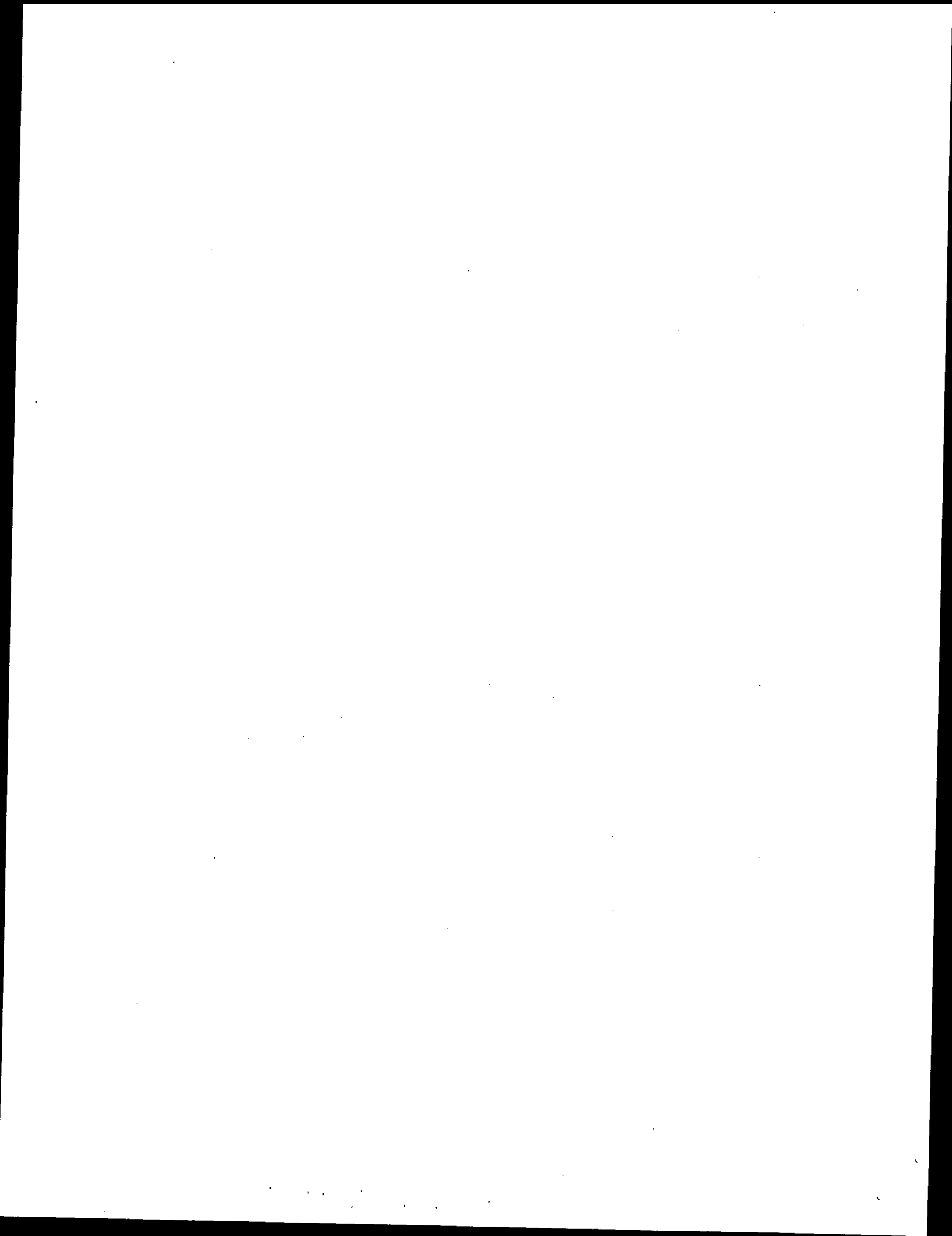
RESULT 13
US-09-505-532-29293/c
; Sequence 29293, Application US/09505532
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Plant Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15478)B
; CURRENT APPLICATION NUMBER: US/09/505,532
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: US/09/505,532
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 51470
; SEQ ID NO 29293
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
US-09-505-532-29293

Query Match      68.0%; Score 20.4; DB 22; Length 599;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTCTTTTGATGGCT 30
    |||||
Db 188 TCCGGTTTCATCGCCACTGTTTGTGATCACT 159

```

Search completed: August 27, 2003, 10:57:21  
Job time : 136.55 secs



Thu Aug 28 09:00:59 2003

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 01:11:55 ; Search time 8.8736 Seconds  
(without alignments)  
3071.976 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30  
Sequence: 1 tcggatccagcgctctgtttgatgct 30

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 780815 seqs, 454324936 residues

Total number of hits satisfying chosen parameters: 1561630

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA New:  
1: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq:  
2: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq:  
3: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:  
4: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:  
5: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:  
6: /cgn2\_6/ptodata/2/pna/US11\_NEW\_COMB.seq:  
7: /cgn2\_6/ptodata/2/pna/US12\_NEW\_COMB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

-Result	No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	6	US-10-349-374-1	Sequence 1, Appli
2	28	93.3	1133	6	US-10-349-374-3	Sequence 3, Appli
3	20.6	68.7	2256646	6	US-10-470-565-1	Sequence 1, Appli
4	19.8	66.0	6645	6	US-10-464-368-44	Sequence 44, Appli
5	19.4	64.7	570	6	US-10-470-888-193	Sequence 193, App
6	19.4	64.7	13794	7	US-60-485-450-11872	Sequence 11872, A
7	18.6	62.0	34288	7	US-60-487-610-19242	Sequence 19242, A
8	18.4	61.3	352	6	US-10-629-771-1755	Sequence 1755, App
9	18	60.0	2380	6	US-10-302-172-785	Sequence 785, App
10	18	60.0	2452	6	US-10-302-172-786	Sequence 786, App
11	18	60.0	27464	7	US-60-487-610-19431	Sequence 19431, A
12	18	60.0	8059021	5	US-09-947-314-53	Sequence 53, Appli
13	17.8	59.3	545	5	US-09-897-516A-1068	Sequence 1068, Ap
14	17.8	59.3	565	6	US-10-357-930-44477	Sequence 44477, A
15	17.8	59.3	1263	6	US-10-293-244-248	Sequence 248, App
16	17.8	59.3	1386	1	PCT-US02-38594-23	Sequence 23, Appli
17	17.8	59.3	1386	1	PCT-US02-38594-27	Sequence 27, Appli
18	17.8	59.3	1386	1	PCT-US02-38594-29	Sequence 29, Appli
19	17.8	59.3	1386	1	PCT-US02-38594-31	Sequence 31, Appli
20	17.8	59.3	1386	1	PCT-US02-38594-33	Sequence 33, Appli
21	17.8	59.3	2147	6	US-10-357-930-29272	Sequence 29272, A
22	17.8	59.3	2174	6	US-10-357-930-23400	Sequence 23400, A
23	17.8	59.3	2174	6	US-10-357-930-29287	Sequence 29287, A
24	17.8	59.3	2223	1	PCT-US02-38594-37	Sequence 37, Appli
25	17.8	59.3	2223	1	PCT-US02-38594-41	Sequence 41, Appli
26	17.8	59.3	2230	1	PCT-US02-38594-43	Sequence 43, Appli

ALIGNMENTS

RESULT 1

US-10-349-374-1

; Sequence 1, Application US/10349374

; GENERAL INFORMATION:

; APPLICANT: Donald L. Durden

; TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES  
ASPARAGINASE IN THE TREATMENT OF HUMAN  
HEMATOLOGIC AND AUTOIMMUNE DISEASE

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; Suite 4700

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM: 3.5" Diskette, 1.44 Mb

; MEDIUM TYPE: storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: FastSEQ for Windows 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/349,374

; FILING DATE: 21-Jan-2003

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/094,435

; FILING DATE: Filed Herewith

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; APPLICATION NUMBER: 60/049,085

; FILING DATE: June 9, 1997

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 234/274

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 30 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

Sequence 25, Appli  
Sequence 3, Appli  
Sequence 21, Appli  
Sequence 13, Appli  
Sequence 39, Appli  
Sequence 45, Appli  
Sequence 17, Appli  
Sequence 15, Appli  
Sequence 19, Appli  
Sequence 1021, Ap  
Sequence 55, Appli  
Sequence 54, Appli  
Sequence 168, App  
Sequence 115, App  
Sequence 3539, Ap  
Sequence 3540, Ap  
Sequence 3541, Ap  
Sequence 3542, Ap  
Sequence 3543, Ap



TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (265)..(501)  
US-10-470-868-193

Query Match 64.7%; Score 19.4; DB 6; Length 570;  
Best Local Similarity 79.3%; Pred. No. 28;  
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 2 CCGATCCAGCGCTCTGTTTGGTGGCT 30  
DB 36 CTGGATCCTGAGGCTCTTTTGTGGCT 8

RESULT 6  
US-60-485-450-11872/c  
; Sequence 11872, Application US/60485450  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001470  
; CURRENT APPLICATION NUMBER: US/60/485,450  
; CURRENT FILING DATE: 2003-07-09  
; NUMBER OF SEQ ID NOS: 47859  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 11872  
; LENGTH: 13794  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-485-450-11872

Query Match 64.7%; Score 19.4; DB 7; Length 13794;  
Best Local Similarity 79.3%; Pred. No. 37;  
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 2 CCGATCCAGCGCTCTGTTTGGTGGCT 30  
DB 6027 CTGGATCCTGAGGCTCTTTTGTGGCT 5999

RESULT 7  
US-60-487-610-19242  
; Sequence 19242, Application US/60487610  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: HUANG, Hongjin  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001469  
; CURRENT APPLICATION NUMBER: US/60/487,610  
; CURRENT FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 97101  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 19242  
; LENGTH: 34288  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-487-610-19242

Query Match 62.0%; Score 18.6; DB 7; Length 34288;  
Best Local Similarity 84.0%; Pred. No. 90;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 4 GGATCCAGCGCTCTGTTTGGTGGCT 28  
DB 33828 GGATCCAGCGCTGAGTTTGTGGCT 33852

RESULT 8  
US-10-629-771-1755/c  
; Sequence 1755, Application US/10629771  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Lewis T.  
; APPLICANT: Escobedo, Jaime  
; APPLICANT: Innis, Michael A.  
; APPLICANT: Garcia, Pablo Dominguez  
; APPLICANT: Klinger, Julie  
; APPLICANT: Reinhard, Christoph  
; APPLICANT: Randazzo, Filippo  
; APPLICANT: Kennedy, Giulia C.  
; APPLICANT: Pot, David  
; APPLICANT: Lamson, George  
; APPLICANT: Drmanac, Radoje  
; APPLICANT: Crkvenjakov, Radomir  
; APPLICANT: Dickson, Mark  
; APPLICANT: Drmanac, Snezana  
; APPLICANT: Labat, Ivan  
; APPLICANT: Leshkowitz, Dena  
; APPLICANT: Kita, David  
; APPLICANT: Garcia, Veronica  
; APPLICANT: Jones, William Lee  
; APPLICANT: Stache-Crain, Birjit  
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
; TITLE OF INVENTION: PRODUCTS  
; FILE REFERENCE: 2300-1598CON  
; CURRENT APPLICATION NUMBER: US/10/629,771  
; CURRENT FILING DATE: 2003-07-28  
; PRIOR APPLICATION NUMBER: 09/611,527  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/142,310  
; PRIOR FILING DATE: 1999-07-02  
; PRIOR APPLICATION NUMBER: 60/142,311  
; PRIOR FILING DATE: 1999-07-02  
; NUMBER OF SEQ ID NOS: 3351  
; SOFTWARE: Fast-SEQ for Windows Version 3.0  
; SEQ ID NO 1755  
; LENGTH: 352  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-629-771-1755

Query Match 61.3%; Score 18.4; DB 6; Length 352;  
Best Local Similarity 78.6%; Pred. No. 73;  
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGATCCAGCGCTCTGTTTGGTGGCT 28  
DB 295 TCCGTTTCCAGGCTCTTGTGTGGTGG 268

RESULT 9  
US-10-302-172-785/c  
; Sequence 785, Application US/10302172  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: Novel Arginine-rich Protein-like Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 803 ICNCP  
; CURRENT APPLICATION NUMBER: US/10/302,172  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/225,251  
; PRIOR FILING DATE: 2002-08-20  
; PRIOR APPLICATION NUMBER: PCT US02/05095  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 09/799,451  
; PRIOR FILING DATE: 2001-03-05  
; NUMBER OF SEQ ID NOS: 950

```
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 785
; LENGTH: 2380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (359)..(1618)
US-10-302-172-785
```

```
Query Match 60.0%; Score 18; DB 6; Length 2380;
Best Local Similarity 80.8%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 TCCGATCCAGCGCCTCTGTTTGTAT 26
||| ||||| ||||| ||||| |||||
Db 1617 TCATCATCCAGCGCCTCTGTTGTAT 1592
```

## RESULT 10

```
US-10-302-172-786/c
; Sequence 786, Application US/10302172
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Arginine-rich Protein-like Nucleic Acids and
; FILE REFERENCE: 803 1CNCp
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 786
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (359)..(1690)
US-10-302-172-786
```

```
Query Match 60.0%; Score 18; DB 6; Length 2452;
Best Local Similarity 80.8%; Pred. No. 1.3e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 TCCGATCCAGCGCCTCTGTTTGTAT 26
||| ||||| ||||| ||||| |||||
Db 1689 TCATCATCCAGCGCCTCTGTTGTAT 1664
```

## RESULT 11

```
US-60-487-610-19431
; Sequence 19431, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19431
; LENGTH: 277464
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(277464)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Table)
US-60-487-610-19431
```

```
Query Match 60.0%; Score 18; DB 7; Length 277464;
Best Local Similarity 80.8%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 2 CCGATCCAGCGCCTCTGTTTGTATG 27
||| ||||| ||||| ||||| |||||
Db 173535 CCGGTCCAGCGCCTCTGTCTGTATG 173560
```

## RESULT 12

```
US-09-947-914-53/c
; Sequence 53, Application US/09947914
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON
; FILE REFERENCE: 21, METHODS OF DETECTION, AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/947,914
; CURRENT FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 75
; SEQ ID NO 53
; LENGTH: 8059021
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8059021)
; OTHER INFORMATION: n = A,T,C or G
US-09-947-914-53
```

```
Query Match 60.0%; Score 18; DB 5; Length 8059021;
Best Local Similarity 80.8%; Pred. No. 1e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 TCCGATCCAGCGCCTCTGTTTGTAT 26
||| ||||| ||||| ||||| |||||
Db 6719885 TCCTGAACCAAGTGCCTGTGTTTCAT 6719860
```

## RESULT 13

```
US-09-897-516A-1068/c
; Sequence 1068, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesling, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 1068
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (260)..(544)
```

OTHER INFORMATION:  
US-09-897-516A-1068  
Query Match 59.3%; Score 17.8; DB 5; Length 545;  
Best Local Similarity 75.9%; Pred. No. 1.4e+02;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 TCGGATCCAGCGCCTCTCTTTTGATGGC 29  
DB 160 TCGGATCCAGCGCCTCTCTTTTGATGGC 132  
RESULT 14  
US-10-357-930-44477  
Sequence 44477, Application US/10357930  
GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Endegre, Wilson  
APPLICANT: Monahan, John  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
HUMAN PROSTATE CANCER  
FILE REFERENCE: MRI-007BCN  
CURRENT APPLICATION NUMBER: US/10/357,930  
CURRENT FILING DATE: 2003-02-04  
PRIOR APPLICATION NUMBER: 09/785,276  
PRIOR FILING DATE: 2003-02-16  
PRIOR APPLICATION NUMBER: 60/183,319  
PRIOR FILING DATE: 2000-02-17  
PRIOR APPLICATION NUMBER: 60/189,862  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: 60/207,454  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: 60/211,314  
PRIOR FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: 60/219,007  
PRIOR FILING DATE: 2000-07-18  
PRIOR APPLICATION NUMBER: 60/255,281  
PRIOR FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 62232  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 44477  
LENGTH: 565  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-357-930-44477  
Query Match 59.3%; Score 17.8; DB 6; Length 1263;  
Best Local Similarity 75.9%; Pred. No. 1.5e+02;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 2 CCGGATCCAGCGCCTCTCTTTTGATGGCT 30  
DB 616 CTGGATCTGGCGCTTCTACTTTGAGGCT 644  
Search completed: August 27, 2003, 11:03:26  
Job time : 25.8736 secs

OTHER INFORMATION:  
US-09-897-516A-1068  
Query Match 59.3%; Score 17.8; DB 5; Length 545;  
Best Local Similarity 75.9%; Pred. No. 1.4e+02;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 TCGGATCCAGCGCCTCTCTTTTGATGGC 29  
DB 160 TCGGATCCAGCGCCTCTCTTTTGATGGC 132  
RESULT 14  
US-10-357-930-44477  
Sequence 44477, Application US/10357930  
GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Endegre, Wilson  
APPLICANT: Monahan, John  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
HUMAN PROSTATE CANCER  
FILE REFERENCE: MRI-007BCN  
CURRENT APPLICATION NUMBER: US/10/357,930  
CURRENT FILING DATE: 2003-02-04  
PRIOR APPLICATION NUMBER: 09/785,276  
PRIOR FILING DATE: 2003-02-16  
PRIOR APPLICATION NUMBER: 60/183,319  
PRIOR FILING DATE: 2000-02-17  
PRIOR APPLICATION NUMBER: 60/189,862  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: 60/207,454  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: 60/211,314  
PRIOR FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: 60/219,007  
PRIOR FILING DATE: 2000-07-18  
PRIOR APPLICATION NUMBER: 60/255,281  
PRIOR FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 62232  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 44477  
LENGTH: 565  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-357-930-44477  
Query Match 59.3%; Score 17.8; DB 6; Length 565;  
Best Local Similarity 75.9%; Pred. No. 1.4e+02;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 2 CCGGATCCAGCGCCTCTCTTTTGATGGCT 30  
DB 372 CTGGATCTGGCGCTTCTACTTTGAGGCT 400  
RESULT 15  
US-10-293-244-248  
Sequence 248, Application US/10293244  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
APPLICANT: Tang, Y. Tom et al  
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
FILE REFERENCE: 21272-029  
CURRENT APPLICATION NUMBER: US/10/293,244  
CURRENT FILING DATE: 2002-11-12  
PRIOR APPLICATION NUMBER: Not Yet Assigned  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: 09/728,422  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: 09/693,325  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 09/663,561

OTHER INFORMATION:  
US-09-897-516A-1068  
Query Match 59.3%; Score 17.8; DB 5; Length 545;  
Best Local Similarity 75.9%; Pred. No. 1.4e+02;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 TCGGATCCAGCGCCTCTCTTTTGATGGC 29  
DB 160 TCGGATCCAGCGCCTCTCTTTTGATGGC 132  
RESULT 14  
US-10-357-930-44477  
Sequence 44477, Application US/10357930  
GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Endegre, Wilson  
APPLICANT: Monahan, John  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
HUMAN PROSTATE CANCER  
FILE REFERENCE: MRI-007BCN  
CURRENT APPLICATION NUMBER: US/10/357,930  
CURRENT FILING DATE: 2003-02-04  
PRIOR APPLICATION NUMBER: 09/785,276  
PRIOR FILING DATE: 2003-02-16  
PRIOR APPLICATION NUMBER: 60/183,319  
PRIOR FILING DATE: 2000-02-17  
PRIOR APPLICATION NUMBER: 60/189,862  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: 60/207,454  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: 60/211,314  
PRIOR FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: 60/219,007  
PRIOR FILING DATE: 2000-07-18  
PRIOR APPLICATION NUMBER: 60/255,281  
PRIOR FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 62232  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 44477  
LENGTH: 565  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-357-930-44477  
Query Match 59.3%; Score 17.8; DB 6; Length 565;  
Best Local Similarity 75.9%; Pred. No. 1.4e+02;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 2 CCGGATCCAGCGCCTCTCTTTTGATGGCT 30  
DB 372 CTGGATCTGGCGCTTCTACTTTGAGGCT 400  
RESULT 15  
US-10-244-248  
Sequence 248, Application US/10293244  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
APPLICANT: Tang, Y. Tom et al  
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
FILE REFERENCE: 21272-029  
CURRENT APPLICATION NUMBER: US/10/293,244  
CURRENT FILING DATE: 2002-11-12  
PRIOR APPLICATION NUMBER: Not Yet Assigned  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: 09/728,422  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: 09/693,325  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 09/663,561



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 27, 2003, 00:51:44 ; Search time 4930.45 Seconds  
(without alignments)  
7529.936 Million cell updates/sec

Title: US-09-937-982-3  
Perfect score: 1133  
Sequence: 1 atgggcagccatcatca.....aagaaagggaattcttctac 1133

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues  
Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				Pending Patents NA Main:*			
1:	/cgn2_6/ptodata/1/pna/PCTUS	COMB.seq.*		1:	/cgn2_6/ptodata/1/pna/PCTUS	COMB.seq.*	
2:	/cgn2_6/ptodata/1/pna/US06	COMB.seq.*		2:	/cgn2_6/ptodata/1/pna/US06	COMB.seq.*	
3:	/cgn2_6/ptodata/1/pna/US07	COMB.seq.*		3:	/cgn2_6/ptodata/1/pna/US07	COMB.seq.*	
4:	/cgn2_6/ptodata/1/pna/US08	COMB.seq.*		4:	/cgn2_6/ptodata/1/pna/US08	COMB.seq.*	
5:	/cgn2_6/ptodata/1/pna/US09	COMB.seq.*		5:	/cgn2_6/ptodata/1/pna/US09	COMB.seq.*	
6:	/cgn2_6/ptodata/1/pna/US10	COMB.seq.*		6:	/cgn2_6/ptodata/1/pna/US10	COMB.seq.*	
7:	/cgn2_6/ptodata/1/pna/US11	COMB.seq.*		7:	/cgn2_6/ptodata/1/pna/US11	COMB.seq.*	
8:	/cgn2_6/ptodata/1/pna/US12	COMB.seq.*		8:	/cgn2_6/ptodata/1/pna/US12	COMB.seq.*	
9:	/cgn2_6/ptodata/1/pna/US13	COMB.seq.*		9:	/cgn2_6/ptodata/1/pna/US13	COMB.seq.*	
10:	/cgn2_6/ptodata/1/pna/US14	COMB.seq.*		10:	/cgn2_6/ptodata/1/pna/US14	COMB.seq.*	
11:	/cgn2_6/ptodata/1/pna/US15	COMB.seq.*		11:	/cgn2_6/ptodata/1/pna/US15	COMB.seq.*	
12:	/cgn2_6/ptodata/1/pna/US16	COMB.seq.*		12:	/cgn2_6/ptodata/1/pna/US16	COMB.seq.*	
13:	/cgn2_6/ptodata/1/pna/US17	COMB.seq.*		13:	/cgn2_6/ptodata/1/pna/US17	COMB.seq.*	
14:	/cgn2_6/ptodata/1/pna/US18	COMB.seq.*		14:	/cgn2_6/ptodata/1/pna/US18	COMB.seq.*	
15:	/cgn2_6/ptodata/1/pna/US19	COMB.seq.*		15:	/cgn2_6/ptodata/1/pna/US19	COMB.seq.*	
16:	/cgn2_6/ptodata/1/pna/US20	COMB.seq.*		16:	/cgn2_6/ptodata/1/pna/US20	COMB.seq.*	
17:	/cgn2_6/ptodata/1/pna/US21	COMB.seq.*		17:	/cgn2_6/ptodata/1/pna/US21	COMB.seq.*	
18:	/cgn2_6/ptodata/1/pna/US22	COMB.seq.*		18:	/cgn2_6/ptodata/1/pna/US22	COMB.seq.*	
19:	/cgn2_6/ptodata/1/pna/US23	COMB.seq.*		19:	/cgn2_6/ptodata/1/pna/US23	COMB.seq.*	
20:	/cgn2_6/ptodata/1/pna/US24	COMB.seq.*		20:	/cgn2_6/ptodata/1/pna/US24	COMB.seq.*	
21:	/cgn2_6/ptodata/1/pna/US25	COMB.seq.*		21:	/cgn2_6/ptodata/1/pna/US25	COMB.seq.*	
22:	/cgn2_6/ptodata/1/pna/US26	COMB.seq.*		22:	/cgn2_6/ptodata/1/pna/US26	COMB.seq.*	
23:	/cgn2_6/ptodata/1/pna/US27	COMB.seq.*		23:	/cgn2_6/ptodata/1/pna/US27	COMB.seq.*	
24:	/cgn2_6/ptodata/1/pna/US28	COMB.seq.*		24:	/cgn2_6/ptodata/1/pna/US28	COMB.seq.*	
25:	/cgn2_6/ptodata/1/pna/US29	COMB.seq.*		25:	/cgn2_6/ptodata/1/pna/US29	COMB.seq.*	
26:	/cgn2_6/ptodata/1/pna/US30	COMB.seq.*		26:	/cgn2_6/ptodata/1/pna/US30	COMB.seq.*	
27:	/cgn2_6/ptodata/1/pna/US31	COMB.seq.*		27:	/cgn2_6/ptodata/1/pna/US31	COMB.seq.*	
28:	/cgn2_6/ptodata/1/pna/US32	COMB.seq.*		28:	/cgn2_6/ptodata/1/pna/US32	COMB.seq.*	
29:	/cgn2_6/ptodata/1/pna/US33	COMB.seq.*		29:	/cgn2_6/ptodata/1/pna/US33	COMB.seq.*	
30:	/cgn2_6/ptodata/1/pna/US34	COMB.seq.*		30:	/cgn2_6/ptodata/1/pna/US34	COMB.seq.*	
31:	/cgn2_6/ptodata/1/pna/US35	COMB.seq.*		31:	/cgn2_6/ptodata/1/pna/US35	COMB.seq.*	
32:	/cgn2_6/ptodata/1/pna/US36	COMB.seq.*		32:	/cgn2_6/ptodata/1/pna/US36	COMB.seq.*	
33:	/cgn2_6/ptodata/1/pna/US37	COMB.seq.*		33:	/cgn2_6/ptodata/1/pna/US37	COMB.seq.*	
34:	/cgn2_6/ptodata/1/pna/US38	COMB.seq.*		34:	/cgn2_6/ptodata/1/pna/US38	COMB.seq.*	
35:	/cgn2_6/ptodata/1/pna/US39	COMB.seq.*		35:	/cgn2_6/ptodata/1/pna/US39	COMB.seq.*	
36:	/cgn2_6/ptodata/1/pna/US40	COMB.seq.*		36:	/cgn2_6/ptodata/1/pna/US40	COMB.seq.*	
37:	/cgn2_6/ptodata/1/pna/US41	COMB.seq.*		37:	/cgn2_6/ptodata/1/pna/US41	COMB.seq.*	
38:	/cgn2_6/ptodata/1/pna/US42	COMB.seq.*		38:	/cgn2_6/ptodata/1/pna/US42	COMB.seq.*	
39:	/cgn2_6/ptodata/1/pna/US43	COMB.seq.*		39:	/cgn2_6/ptodata/1/pna/US43	COMB.seq.*	
40:	/cgn2_6/ptodata/1/pna/US44	COMB.seq.*		40:	/cgn2_6/ptodata/1/pna/US44	COMB.seq.*	
41:	/cgn2_6/ptodata/1/pna/US45	COMB.seq.*		41:	/cgn2_6/ptodata/1/pna/US45	COMB.seq.*	
42:	/cgn2_6/ptodata/1/pna/US46	COMB.seq.*		42:	/cgn2_6/ptodata/1/pna/US46	COMB.seq.*	
43:	/cgn2_6/ptodata/1/pna/US47	COMB.seq.*		43:	/cgn2_6/ptodata/1/pna/US47	COMB.seq.*	

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1133	100.0	1133	39	US-09-937-982-3
2	1129.8	99.7	1133	33	Sequence 3, Appli
C 3	302.8	26.7	9082	37	Sequence 1293, Ap
C 4	302.8	26.7	9082	37	Sequence 1294, Ap



RESULT 2

US-09-773-260-3

; Sequence 3, Application US/09773260

; GENERAL INFORMATION:

; APPLICANT: Donald L Durden

; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES

; ASPARAGINASE IN THE TREATMENT OF HUMAN

; HEMATOLOGIC AND AUTOIMMUNE DISEASE

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Lyon & Lyon

; STREET: 633 West Fifth Street

; Suite 4700

; CITY: Los Angeles

; STATE: California

; COUNTRY: U.S.A.

; ZIP: 90071-2066

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS 5.0

; SOFTWARE: FastSeq for Windows 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/773,260

; FILING DATE: 31-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/094,435

; FILING DATE: 1998-06-09

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/POCKET NUMBER: 234/274

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1133 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Query Match 99.7%; Score 1129.8; DB 33; Length 1133;

Best Local Similarity 99.8%; Pred No. 0;

Matches 1131; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGGACAGCCATCATCATCATATAGCAGCGCTGGTGGCGCGGAGCCAT 60

Db 1 ATGGGACAGCCATCATCATCATATAGCAGCGCTGGTGGCGCGGAGCCAT 60

Qy 61 ATGGCTAGCATGCTGGTGGACAGCAATGGTCCGGATCCAGCGCTCTGTTTGATG 120

Db 61 ATGGCTAGCATGCTGGTGGACAGCAATGGTCCGGATCCAGCGCTCTGTTTGATG 120

Qy 121 GCTAAACCCCAAGTACTCTAGCCACAGGAGCCATCGCTGGTTCGGGGGAATCT 180

Db 121 GCTAAACCCCAAGTACTCTAGCCACAGGAGCCATCGCTGGTTCGGGGGAATCT 180

Qy 181 AGCGTCAAGTAGTACTCTGCTGGAGCAGTCAACCGTTGTAAGCTTCTTCAGCCGTC 240

Db 181 AGCGTCAAGTAGTACTCTGCTGGAGCAGTCAACCGTTGTAAGCTTCTTCAGCCGTC 240

Qy 241 CCTGCCATCAAGCAGCTAGCCACCATCAAGGTGACAGATCTCAAGATTGGTCCCA 300

Db 241 CCTGCCATCAAGCAGCTAGCCACCATCAAGGTGACAGATCTCAAGATTGGTCCCA 300

241 CCTGCCATCAAGCAGCTAGCCACCATCAAGGTGAAAGATCTCAAGCATTTGGTCCCAA 300

Qy 301 GAGATGACGGTAAAGTGTGGCTTAACTAGCCAAAGCGTGTCAATAGTCTCTCGCCCAA 360

Db 301 GAGATGACGGTAAAGTGTGGCTTAACTAGCCAAAGCGTGTCAATAGTCTCTCGCCCAA 360

Qy 361 AAAGAGACCGAAGCGGTGATCATCACCCTAGAACTGACCATCGAAGAGACCGCTTTC 420

Db 361 AAAGAGACCGAAGCGGTGATCATCACCCTAGAACTGACCATCGAAGAGACCGCTTTC 420

Qy 421 TTCTCAACCTCAACGCTGAAAGGCAAAACCTGTCTGCTGTAGGGGCCATGCGTCCA 480

Db 421 TTCTCAACCTCAACGCTGAAAGGCAAAACCTGTCTGCTGTAGGGGCCATGCGTCCA 480

Qy 481 GGCTCTTCCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540

Db 481 GGCTCTTCCATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540

Qy 541 AACAAGCCTCTACTAAAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600

Db 541 AACAAGCCTCTACTAAAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600

Qy 601 AGAAGAGCGACCAAGCTCAACACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAGT 660

Db 601 AGAAGAGCGACCAAGCTCAACACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAGT 660

Qy 661 ABAATCGGCACAGTCTATTATGCAAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720

Db 661 ABAATCGGCACAGTCTATTATGCAAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720

Qy 721 ACCCTTGCAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 780

Db 721 ACCCTTGCAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 780

Qy 781 TAGCTCACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840

Db 781 TAGCTCACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840

Qy 841 GGAATCATCCATCAGGCGATGGGCAATGGGCAATGGGCAATGGGCAATGGGCAATGGGCA 900

Db 841 GGAATCATCCATCAGGCGATGGGCAATGGGCAATGGGCAATGGGCAATGGGCAATGGGCA 900

Qy 901 GAAAAAGCAGCCAAATCAGGCGTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960

Db 901 GAAAAAGCAGCCAAATCAGGCGTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960

Qy 961 ACCACCCAGAGGCTGAAGTGGATGATAAGAACTTGGTTTGGTGGTGGTGGTGGTGGTGGT 1020

Db 961 ACCACCCAGAGGCTGAAGTGGATGATAAGAACTTGGTTTGGTGGTGGTGGTGGTGGTGGT 1020

Qy 1021 AACCTTCAAAAAGCCAGAGTGGTCTTATGTTAGCCCTCACCACCAACTAGTGTATAGAG 1080

Db 1021 AACCTTCAAAAAGCCAGAGTGGTCTTATGTTAGCCCTCACCACCAACTAGTGTATAGAG 1080

Qy 1081 GCGATCCAAAAGATCTTCTCCACCTATTAATCCAAAGAAAGGAATCTCTTCCAC 1133

Db 1081 GCGATCCAAAAGATCTTCTCCACCTATTAATCCAAAGAAAGGAATCTCTTCCAC 1133

RESULT 3

US-09-897-516-1293/c

; Sequence 1293, Application US/09897516

; GENERAL INFORMATION:

; APPLICANT: Corbin, David R.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Huesing, Joseph E.

; APPLICANT: Krasomil-Osterfeld, Karina C.

; APPLICANT: Malvar, Thomas M.

; APPLICANT: Slater, Steven C.

; APPLICANT: Spiridonov, Sergei

; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof

; FILE REFERENCE: 38-21(51847)B

CURRENT APPLICATION NUMBER: US/09/897,516  
CURRENT FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/215, 161  
PRIOR FILING DATE: 2000-06-30  
NUMBER OF SEQ ID NOS: 8409  
SEQ ID NO 1293  
LENGTH: 9082  
TYPE: DNA  
ORGANISM: Xenorhabdus sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3968)..(5734)  
US-09-897-516-1293

Query Match 26.7%; Score 302.8; DB 37; Length 9082;  
Best Local Similarity 58.1%; Pred. No. 3.4e-81;  
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCCAAGTGACTATCTAGCCACAGGAGGACCATCGCTGGTTCGGGGGAATCTAGCGTC 186  
DB 3755 CCAACATTACTGTTCTGGCAACGGTGCACAAATCGCTGGTGGTGAATCTGCAACT 3696

QY 187 AAGAGTAGTACTCTGTGGAGCAGTCAACCGTTGATAGCTTCTTGACCGCTCCCTGCC 246  
DB 3695 AAATCCAGCTACACTGCGGTCAAGTCCGGTGTGATTCATTCGTAATCTGTTCTGCT 3636

QY 247 ATCAACGACCTAGCCACCATCAAGGTTGAACAGTCTCAAGCAATCGCTCCCAAGAGATG 306  
DB 3635 ATCAAAAACATTCGTAATCTGAAGGTGAGCAGGTTGTAGCATCGGTTCTCAGACATG 3576

QY 307 ACAGGTAGGTGTGGCTTAAACTAGCAAGCGTGTCAATGAGTCTCTCGCCCAAAAGAG 366  
DB 3575 AACGATCAGTCTGGCTGACTCTGGGAAAAAATCAATG-----CAGATTGTGATAAA 3522

QY 367 ACCGAAGCGGTGATCATCCCATGGAACATGACCATGGAAGAGACCGCTTCTTCCTC 426  
DB 3521 ACTGACGGTTTGTCTATCCACCGGTACAGATACCATGGAAGAACTGCTTATTTCTC 3462

QY 427 AACCTCAGGTGAAAGCAAAACCTGCTGCTGCTTGTAGGCGCATGCGTCCAGGCTCT 486  
DB 3461 GATCTGACCACTCAGTGTCTAATCACTATCGTAATGTTGGCGCAATGCTCCATCAACG 3402

QY 487 TCCATGAGTGTGATGCGCCCATGATCTCTATAACCGCGTGAATGTAGCGATCAACAAA 546  
DB 3401 GCTCTGGGTGCTGATGGCCCATTTGAACCTCTATAATGCTGTGTGTAGCGTCAGATAA 3342

QY 127 CCCCAAGTGACTATCTAGCCACAGGAGGACCATCGCTGGTTCGGGGGAATCTAGCGTC 186  
DB 3755 CCAACATTACTGTTCTGGCAACGGTGCACAAATCGCTGGTGGTGAATCTGCAACT 3696

QY 187 AAGAGTAGTACTCTGTGGAGCAGTCAACCGTTGATAGCTTCTTGACCGCTCCCTGCC 246  
DB 3695 AAATCCAGCTACACTGCGGGTCAAGTCCGGTGTGATTCATTCGTAATCTGTTCTGCT 3636

QY 247 ATCAACGACCTAGCCACCATCAAGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATG 306  
DB 3635 ATCAAAACATTGCTAATCTGAAGGTGAGCAGGTTGTAGATCGGTTCTCAGGACATG 3576

QY 307 ACAGGTAGGTGTGGCTTAAACTAGCCAGCGTGTCAATGAGTCTCTCGCCCAAAAGAG 366  
DB 3575 AACGATCAGGTTCTGGTACTCTGGCGAAAAAATCAATG-----CAGATTGTGATAAA 3522

QY 367 ACCGAAGCGGTGATCATCCCATGGAACATGACCATGGAAGAGACCGCTTCTTCCTC 426  
DB 3521 ACTGACGGTTTGTCTATCCACCGGTAAAGATCAATGGAAGAACTGCTTATTTCTC 3462

QY 427 AACCTCAGGTGAAAGCAAAACCTGCTGCTGCTTGTAGGCGCATGCGTCCAGGCTCT 486  
DB 3461 GATCTGACCACTCAGTGTCTAATCACTATCGTAATGTTGGCGCAATGCTCCATCAACG 3402

QY 487 TCCATGAGTGTGATGCGCCCATGATCTCTATAACCGCGTGAATGTAGCGATCAACAAA 546  
DB 3401 GCTCTGGGTGCTGATGGCCCATTTGAACCTCTATAATGCTGTGTGTAGCGTCAGATAA 3342

QY 547 GCCTCTACTAACAAAGAGTGGTGTGATGTGATGACGATGAGATTCAACCGCCGACGAA 606  
DB 3341 GACTCTGAAACCGGTGTTTACTTGTATGAATGATTCGTAAATTCATGACGTGAC 3282

QY 607 CGGACCAAGCTCAACACACCGCAGTCAATGCTTTGCTTCGCCCAACACAGGTAAATC 666  
DB 3281 ATCGGTAACCTGACACAACTGAAGTTCAGGCTTCCAGGCAATGACGAGCCACAA 3222

QY 667 GGCACAGTCTAATATGGAAGTTCAGTATTTCACTCAATCCGTTGCACTCAACCCCT 726  
DB 3221 GGCTTCATTCAACCGGTAAGTAACTACTACTCT--GCAGCACAGTGAAGCAGAT 3165

QY 727 GCAAGTCAAGTTGATTTAGCAAAATCGAAGAACTCCCGAGAGTGCATATTTTACGCT 786  
DB 3164 AAAGCTGTTTGTATGTCAGCAAACTGACTGAACCTGCCAAAGTAGGTATTTGTATAAC 3105

QY 787 CACCCCGATGATCTGATGTTTTAGTCAATGAGCGCTTTGAGCAGGAGCCAAAGGAATC 846  
DB 3104 TACTCTAATGCTTCGATCTGCGGTAAAGCATTTGTAGAGAAATGATTATAAGGTATT 3045

QY 847 ATCCATGAGGATGAGGCAATGGGAACCCCTTTCCCTTTGACTCAAAATGCTTTGAAAA 906  
DB 3044 ATCAGTGTGTTGTTGTAACCGGTAACTCTATAAGTCTATTCTGAGCTCTCTGTCAAA 2985

QY 907 GCAGCCCAATCAGGCGTGTGCTGCTCGAAGCTCTAGAGTGGGAGTGGTCCACCAAC 966  
DB 2984 CGGCTAAAGATGGTGTGTTGTTGTTCTTCCACCGTATTCTTCTGCTTCACTACT 2925

QY 547 GCTCTACTAACAAGAGTGGTGAATGATGAACGATGAGATTACGGCGCCAGAGAA 606  
 Db 3341 GACTCTGCAACACCGTGTGTTTACTTGTCTATGATGATTTCTGTAATTCATGGAGGTGAC 3282  
 QY 607 GCGACCAAGCTCAACACCAACCGGAGTCAATGCTATTTGCTTGGCCCAACACAGGTAAATC 666  
 Db 3281 ATCGGTAAATGAGCACTGACCACTGAACTGAGTTCAGGCTTCCAGGAGTGAACGAGGCCACAA 3222  
 QY 667 GGCACAGTCTATTATGCGAAAGTCGAGTATTTTCACTCAATCCGTTGACCTCAACCCCTT 726  
 Db 3221 GGCTTCATTCATPACCGGTAAAGTAACTACTCTCT---GCAGCACGAGTGAAGCAGAT 3165  
 QY 727 GCAAGTGAGTTTGATATGACAAAATCGAAGAACTCCCAAGAGTCGATATTTCTTTACGCT 786  
 Db 3164 AAAGCTGTTTGTGATGTCAGCAACTGACTGAATCGCAAAAGTGTAGATTTGTGTATAC 3105  
 QY 787 CACCCCGATGATGATGATGTTTGTAGTCAATGAGCCCTTTCAGGAGGAGCAAGGAATC 846  
 Db 3104 TACTCTAATGCTTCGATCTGCCGTTAAAGCAATTTGTAGAGATGATATATAAGGTATT 3045  
 QY 847 ATCCATGACGAGTGGCAATGGAAACCCCTTTTCCCTTGTGACTCAAAATGCTCTTGAATA 906  
 Db 3044 ATCAGTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 2985  
 QY 907 GCAGCAATCAGGCGTAGTCTGCTCGTCTGAAAGCTCTAGAGTGGCGAGTGTCCACACC 966  
 Db 2984 GCGGTAAAGATGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2925  
 QY 967 CAAGAGCTGAAGTGTGATGAAGAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1026  
 Db 2924 CAATGCTGAAGTTAAGACAGCAATACGGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2865  
 QY 1027 CAAAAGCAGAGTCTTCTTATGTTAGCCCTCACCAAACTAGTGTATAGAGAGGCGATC 1086  
 Db 2864 CAGAAAGCTGCTGTTCTTCTGCAATGGCTCTGACTCAGACTGCTGACACAGCAAAATTT 2805  
 QY 1087 CAAAAGATCTTCTCACCTATTATTCAGAA 1118  
 Db 2804 CAGAAATGTTTCTTAAATATTAAATTCCTGAA 2773

RESULT 5

US-09-897-516-1295/c  
 ; Sequence 1295, Application US/09897516  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Corbin, David R.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Huesing, Joseph E.  
 ; APPLICANT: Krasomil-Osterfeld, Karina C.  
 ; APPLICANT: Malvar, Thomas M.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Spiridonov, Sergei  
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
 ; FILE REFERENCE: 38-21(51847)B  
 ; CURRENT APPLICATION NUMBER: US/09/897,516  
 ; CURRENT FILING DATE: 2001-06-29  
 ; PRIOR APPLICATION NUMBER: US 60/215, 161  
 ; PRIOR FILING DATE: 2000-06-30  
 ; NUMBER OF SEQ ID NOS: 8409  
 ; SEQ ID NO 1295  
 ; LENGTH: 9082  
 ; TYPE: DNA  
 ; ORGANISM: Xenorhabdus sp.  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (6985) ..(9081)  
 US-09-897-516-1295

Query Match 26.7%; Score 302.8; DB 37; Length 9082;  
 Best Local Similarity 58.1%; Pred. No. 3.4e-81;  
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

RESULT 6

US-09-897-516-1296  
 ; Sequence 1296, Application US/09897516

QY 127 CCCAAGTGAATCTATCTAGCCACAGGAGGACCATCGCTGGTTCGGGGGAATCTAGCGTC 186  
 Db 3755 CCAAAACATTAATCTGTTCTGCAACACGGTGGCAAACTCGCTGGTGGTGAATCTGCAACT 3696  
 QY 187 AAGAGTAGCTACTCTGCTGGAGAGTCAACCGTTGATAAGCTTCTTCAGCGCGTCCCGCC 246  
 Db 3695 AAATCCAGCTACACTGCGGGTCAAGTCGGTGTGATTCATTGCTGAATGCTGTCTCTGCT 3636  
 QY 247 ATCAACGACCTAGCCACCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATG 306  
 Db 3635 ATCAAAACATTAATCTGTAAGTGTGAGAGTGTGTTAGCATCGTCTCAGGACATG 3576  
 QY 307 ACGGTGAAGTGTGGCTTAAACTGACCAAGCGTGTCAATAGCTCTCTCCGCCCAAAAGAG 366  
 Db 3575 AACGATCAGGCTCGCTGACTCTGGCGAAAAAATCAATG-----CAGATTGTGATAAA 3522  
 QY 367 ACCGAACGCGTGAATCATCACCCATGGAATGACACCATGGAGAGACCGCTTCTTCCTC 426  
 Db 3521 ACTGACGGTGTGTCATCACCCAGGTACATACCATGGAAGAAATGCTTATTTCTCTC 3462  
 QY 427 AACCTACGGTGAAGAACCAAAAACTGTCTGCTCTGTAGGCGCCATGCGTCCAGGCTCT 486  
 Db 3461 GATCTGACCACTCAGTGTCTATAAACCTATCGTAATGTTGGCGCAATGGTCCATCAACG 3402  
 QY 487 TCCATGAGTCTGATGCGCCCATGAATCTCTAAGCGCGTGAATGTAGCGATCAACAAA 546  
 Db 3401 GCTCTGGGTGCTGATGCGCCCATTTGAACCTCTATAATGCTTGTGTTAGCGTCAGATAA 3342  
 QY 547 GCCTCTACTAAACAAAGAGTGGTGTATGTGATGAACAGTGAATTCACGCCGCCAGAGAA 606  
 Db 3341 GACTCTGCAACCGTGTGTTTACTTGTCTAATGATGATTCGTAAATTCATGGACGTGAC 3282  
 QY 607 GCGACCAAGTCAACCAACCGCAGTCAATGCTTGTCTGCCCAACACACAGGTAAAAATC 666  
 Db 3281 ATCGTAAACTGAGCAAACTGAAATTCAGGCTTTCAGGCGAGTGAACGAGGCCACAA 3222  
 QY 667 GGCACAGTCTATTATGCAAGTCAAGTATTTCACTCAATCGTTCGAGCTCACAACCTTT 726  
 Db 3221 GGCTTCATTCATACGTAAGTAACTACTACTCT---GCAGACCAAGTGAAGCAGAT 3165  
 QY 727 GCAAGTGAGTTGATATTAAGCAAAATCGAAAGTCTCCCGAGAGTCGATATTTCTTACGCT 786  
 Db 3164 AAAGCTGTTTGTGATGTCAGCAAACTGACTGAACTGCCAAAGTAGTATTGTGTATAC 3105  
 QY 787 CACCCGATGATGATGATGTTTGTAGTCAATGCGCCCTTCAGGCGAGGCGCAAGGAATC 846  
 Db 3104 TACTCTAATGCTTCGATCTGCCGTTAAAGCATTTGTAGAGATGATATATAAGGTATT 3045  
 QY 847 ATCCATGACGAGTGGCAATGGAAACCCCTTTCCTTTGACTCAAAATGCTCTTGAATAA 906  
 Db 3044 ATCAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2985  
 QY 907 GCGCAAAATCAGGCGTAGTCTGCTCGTCAAGTCTTAGAGTGGGAGTGTCCACACC 966  
 Db 2984 GCGGTAAAGATGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2925  
 QY 967 CAAGAGCTGAAGTGAATGAAGAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1026  
 Db 2924 CAAAATGCTGAAGTTAATGACAGCAAAATACGGTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 2865  
 QY 1027 CAAAAGCAGAGTCTTCTTATGTTAGCCCTCACCAAACTAGTGTATAGAGAGGCGATC 1086  
 Db 2864 CAGAAAGCTGCTGTTCTTCTGCAATGGCTCTGACTCAGACTGCTGACACAGCAAAATTT 2805  
 QY 1087 CAAAAGATCTTCTCACCTATTATTCAGAA 1118  
 Db 2804 CAGAAATGTTTCTTAAATATTAAATTCCTGAA 2773



QY 367 ACCGAGCGGTGATCATCACCCATGGAATGACACATGGAAGAGACCGCTTTCTTCCTC 426  
 Db ACTGACGGTTTTGTTCATCACCCACGGTACAGATACCATGGAAGAACTGCTTATTTCTTC 5621  
 QY 427 AACCTCAGGTGAAAGCCAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486  
 Db 5622 GATCTGACCACTAGTGTCATAAACCTTATGTAATGCTGTTGTTGAGGTCAGATAA 5681  
 QY 487 TCCATGAGTGTGATGGCCCATGAATCTCTATACCGCGTGAATGTAAGCATCAACAAA 546  
 Db 5682 GCTCTGGGTGCTGATGGCCCATGAACCTCTATATGCTGTTGTTGAGGTCAGATAA 5741  
 QY 547 GCCTCTACTAAACAAAGGAGTGGTGAATGTAAGACGATGAGATTCACGCCGCCAGAAA 606  
 Db 5742 GACTCTGCAAAACCGGTGTTTACTTGTCTATGAATGATCTGTGAATTCATGACGCTG 5801  
 QY 607 GCGACCAAGCTCAACACACCGCAGTCAATGCTTCTGCTTCCAGGAGTGAACGAGGCA 666  
 Db 5802 ATCCGTAACTTGAGCACAACTGAAGTTCAGGCTTCCAGGAGTGAACGAGGCAACAA 5861  
 QY 667 GGCACAGTCTATTATGGCAAAAGTCGAGTATTTCACTCAATCCGTTTCGACCTCACACCT 726  
 Db 5862 GCTCTCATTCATACCGGTAAAGTAACTACTACTCT--GCAGCACCGAGTGAAGCAGAT 5918  
 QY 727 GCAAGTGAATTTGATATAGCAAAATCGAAGAACTCCCGCAGAGTGAATTTCTTACGCT 786  
 Db 5919 AAAGCTGTTTTGATGTGAGCAAACTGACTGAACCTGCCAAAGTAGTATTTGTATTAAC 5978  
 QY 787 CACCCGATGATCTGATGTTTTAGTCAATCGAGCCCTTCAGGAGGAGCCCAAGGAATC 846  
 Db 5979 TACTCTAATGCTTCGATCTGCGGTAAAGCAATTTAGAGAAATGATTAAGAGTATTT 6038  
 QY 847 ATCCATCAGGAGTGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTCTTGAAAAA 906  
 Db 6039 ATCAGTCTGCTGTTGGTAAACGTAACATCTATAAGTCTATTCTGAGCTCTCTGTACA 6098  
 QY 907 GCAGCCAAATCAGGCGTATGCTGCTCGCTCGAAGCTCTAGAGTGGGAGTGGTTCCACCC 966  
 Db 6099 GCGGTAAAGATGGTGTGTTGTTGCTTCCAAACCGTATTTCTTTCGTTTCTACTACT 6158  
 QY 967 CAAGAGCTGAAGTGAATGAAGAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1026  
 Db 6159 CAAATGCTGAAGTAAATGACAGCAAAATACGCTTTGTTGTTGTTGTTGTTGTTGTTGTT 6218  
 QY 1027 CAAAAAGCCAGAGTCTTTCTTATGTTAGCCCTCAACAAATAGTATGATAGAGGCGATC 1086  
 Db 6219 CAGAAAGCTCGTGTCTTCTGCAATGGCTGACTCAGACTGCTGACATGCTGACACGAAA 6278  
 QY 1087 CAAAAAGATCTTCTCCACCTATTAAATCCAAAGAA 1118  
 Db 6279 CAGAAATGTTTTCTAAATATTAAATCTCTGAA 6310

RESULT 8

US-09-897-516-1298  
 ; Sequence 1298, Application US/09897516  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Corbin, David R.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Huesing, Joseph E.  
 ; APPLICANT: Krasomil-Osterfeld, Karina C.  
 ; APPLICANT: Malvar, Thomas M.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Spiridonov, Sergei  
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
 ; FILE REFERENCE: 38-21(51847)B  
 ; CURRENT APPLICATION NUMBER: US/09/897,516  
 ; CURRENT FILING DATE: 2001-06-29  
 ; PRIOR APPLICATION NUMBER: US 60/215, 161  
 ; PRIOR FILING DATE: 2000-06-30  
 ; NUMBER OF SEQ ID NOS: 8409  
 ; SEQ ID NO 1298

; LENGTH: 9082  
 ; TYPE: DNA  
 ; ORGANISM: Xenorhabdus sp.  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (5259)..(6302)  
 ; US-09-897-516-1298  
 Query Match 26.7%; Score 302.8; DB 37; Length 9082;  
 Best Local Similarity 58.1%; Pred. No. 3.4e-81;  
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;  
 QY 127 CCCCAAGTGAATCTCTAGCCACAGAGGACCATCGCTGGTTCGGGGAACTTAGCGCTC 186  
 Db 5328 CCAAAACATTACTGTTCTGGCAACGGGTGGCACAATCGCTGGTGGTGAATCTGCAACT 5387  
 QY 187 AAGAGTAGCTACTCTCTGGAGCAGTCAACGTTGATAAGCTTCTTGCAGCGCTCCCTGCC 246  
 Db 5388 AATCCAGCTACACTCGGGTCAAGTGGTGTGATTCATTTGCTGAATGCTGTCTCTGCT 5447  
 QY 247 ATCAACGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATG 306  
 Db 5448 ATCAAAAACATTGCTTAATCTGAAAGGTGAGCAGGTTCTTAGCATCGGTTCTCAGGACATG 5507  
 QY 307 ACGGTGAAGTGTGGCTTAAACTAGCAACGGTGTCAATGAGCTCTCGCCCAAAAAGAG 366  
 Db 5508 AACGATCAGGCTCTGGTGAATCTTGGCGAAAAAATCAATG-----CAGATTGTGATAAA 5561  
 QY 367 ACCGAAGCGTGAATCATCAACCATGGAACGACACCATGGAAGAGACCGCTTTCTTCCTC 426  
 Db 5562 ACTGACGTTTTGCTATCACCACGGTACAGATACCATGGAAGAACTGCTTATTTCTCTC 5621  
 QY 427 AACCTCAGCGTGAAGCCAAACCTGCTGCTTGTAGCGCCATGCGTCCAGGCTCT 486  
 Db 5622 GATCTGACCACTCAGTGTATTAACCTATGTAATGTTGGCGCAATGCGTCCATCAAG 5681  
 QY 487 TCCATGAGTGTGATGGCCCATGAATCTCTATAACCGCGTGAATGTAAGCATCAACAAA 546  
 Db 5682 GCTCTGGTGTCTGATGGCCCATTGAACCTCTATAATGCTGTTGTTGAGCGTCAGATAAA 5741  
 QY 547 GCCTCTACTAAACAAAGGAGTGGTGAATGATGAACATGATGATTCAGCGCCAGAGAA 606  
 Db 5742 GACTCTGCAAAACCGTGGTGTGTTTACTTGTCTATGAATGATTTCTGTAATTCATGAGCGTGA 5801  
 QY 607 GCAGCAAGCTCAACACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAGGTAAATATC 666  
 Db 5802 ATCGGTAACTGAGCACCACTGAAGTTCAGGCTTCCAGGCACTGAACGAGGCCCAACAA 5861  
 QY 667 GGCACAGTCTATTATGGCAAAAGTGAATTTCACTCAATCCGTTTCGACCTCACACCTTT 726  
 Db 5862 GGCTTCATTCATAACCGTAAAGTAACTACTACTCT---GCAGCACCGAGTGAAGCAGAT 5918  
 QY 727 GCAAGTGAATTTGATATAGCAAAATCGAAGAACTCCCGCAGAGTGAATTTCTTACGCT 786  
 Db 5919 AAAGCTGTTTTGATGTGACAAACCTGACTGAACCTGCCAAAGTAGTATTTGTGTAATAC 5978  
 QY 787 CACCCCGATGATCTGATGTTTGTAGTCAATGCAACCTTTCAGGAGGAGCCAAAGGAATC 846  
 Db 5979 TACTCTAATGCTTCCGATCTGCGGTAAAGCAATTTAGAGAAATGATTAAGAGTATTT 6038  
 QY 847 ATCCATGCAAGCAATGGGCAATGGAAACCTTTCCCTTTGACTCAAAATGCTCTTGAAAAA 906  
 Db 6039 ATCAGTCTGCTGTTGGTAAACGTAACATCTATAAGTCTATTCTGAGCTCTCTGTACA 6098  
 QY 907 GCAGCCAAATCAGGCGTATGCTGCTCGCTCGAAGCTCTAGAGTGGGAGTGGTTCCACCC 966  
 Db 6099 GCGGTAAAGATGGTGTGTTGTTGCTTCCAAACCGTATTTCTTTCGTTTCTACTACT 6158  
 QY 967 CAAGAGCTGAAGTGAATGAAGAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1026  
 Db 6159 CAAATGCTGAAGTAAATGACAGCAAAATACGCTTTGTTGTTGTTGTTGTTGTTGTTGTT 6218  
 QY 1027 CAAAAAGCCAGAGTCTTTCTTATGTTAGCCCTCAACAAATAGTATGATAGAGGCGATC 1086  
 Db 6219 CAGAAAGCTCGTGTCTTCTGCAATGGCTGACTCAGACTGCTGACATGCTGACACGAAA 6278  
 QY 1087 CAAAAAGATCTTCTCCACCTATTAAATCCAAAGAA 1118  
 Db 6279 CAGAAATGTTTTCTAAATATTAAATCTCTGAA 6310

Db 6219 CAGAAAGCTGCTGTTCTTCTGCAATGGCTGCTGACTCAGACTGCTGACACAGCAAAAT 6278  
QY 1087 CAAAAGATCTCTCCACCTATTATCAAGAA 1118  
Db 6279 CAGAAATGTTTCTAAATATTAATCTCTGAA 6310

## RESULT 9

US-60-215-161-1293/c  
; Sequence 1293, Application US/60215161  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Huesing, Joseph E.  
; APPLICANT: Krasomil-Osterfeld, Karina C.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)A  
; CURRENT APPLICATION NUMBER: US/60/215,161  
; CURRENT FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 1293  
; TYPE: DNA  
; ORGANISM: Xenorhabdus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3968)..(5734)  
US-60-215-161-1293

Query Match 26.7%; Score 302.8; DB 75; Length 9082;  
Best Local Similarity 58.1%; Pred. No. 3.4e-81;  
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;  
QY 127 CCCAAGTACTATCTAGCCACAGGAGGACCATCTGCTCGGGGAATCTAGCGTC 186  
Db 3755 CCAACATTACTGTTCTGGCAACGGGTGGCACATCGTGGTGGTGAATCTGCACT 3696  
QY 187 AGAGTAGTACTCTGCTGAGCAGTCCCGTTGATGAAGTCTTCTGAGCGGCTCCGCC 246  
Db 3695 AATCCAGCTACACTGCGGGTCAAGTCGGTGTGATTCATTCGTTGATTCGTTCTGCT 3636  
QY 247 ATCAAGCTAGTCTGCTGAGGAGGACAGATCTCAAGCTCTCTCGGCCCAAAAGAG 306  
Db 3635 ATCAAAACATCTGTAATCTGAAGGTGACGAGTCTTGTAGCATCGGTTCTCAGGACATG 3576  
QY 307 ACGGTAAGTGTGGCTTAACTAGCCAGCGTGTCAATGAGTCTCTCGGCCCAAAAGAG 366  
Db 3575 AACGATCAGTCTGGCTGACTCTGGCGAAAAAATCAATG-----CAGATTGTGATAA 3522  
QY 367 ACCGAAGCGTATCATACCCATGGAATGACACCATGGAAGAGACCGCTTTCTCCTC 426  
Db 3521 ACTGACGGTTTGTCTATCACCACCGGTACAGATACCATGGAAGAACTGTTATTCCTC 3462  
QY 427 AACTCAGCGTGAAGAGCAAAACCTGCTGCTGCTAGCGCCATCGCTCCAGGCTCT 486  
Db 3461 GATCTGACCACTCAGTGTCTAATACCTATGATGGTTGGCGAATGCGTCCATCAAG 3402  
QY 487 TCCATGAGTCTGATGGCCCCCATGAATCTCTATAACCGCGTGAATGATGAGGATCAAAA 546  
Db 3401 GCTCTGGTGTGATGGCCCCATTGAACCTCTATAATGCTGTGTGTGATGCTGATAA 3342  
QY 547 GCTCTACTAAACAAAGAGTGTGATGATGAACATGAGATTACGCGGCCAGAGAA 606  
Db 3341 GACTCTGCAACCGTGGTGTGTTTACTTGTGATGAATGATTCGTTAATTCATGACGCTGAC 3282  
QY 607 GCGACCAAGCTCAACACCGAGTCAATGCAATTTGCTTCGCCCAACACAGAGTAAATC 666  
Db 3281 ATCGGTAATCTGACCACTGAATTCAGGCTTTCCAGGAGTGAACGAGGCCCAAA 3222

QY 667 GGCACAGCTATTATGGCAAGTCGAGTATTTTCACTCAATCCGTTCCGACCTCACACCTT 726  
Db 3221 GGTTCATTCATAACGGTAAGTAACACTACTCT--GCAGCACCATGTAAGCAGAT 3165  
QY 727 GCAAGTAGTGTGATATTAGCAAAATCGAAGAACCTCCAGAGTCTGATTTCTTTAGGCT 786  
Db 3164 AAAGCTGTTTGTGATGTCAGCAAACTGACTGAACCTGCAAAAGTAGTATTGTGTATAAC 3105  
QY 787 CACCCGATGATGATGATGTTTGTAGTCAATGAGCCCTTTCAGCAGGAGCCCAAGGATC 846  
Db 3104 TACTCTAATGCTTCCGATCTGCCGGTTAAGCATTTGTAGAGAAATGATTATAAAGGTATT 3045  
QY 847 ATCCATGAGGATGAGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTCTTGA AAAA 906  
Db 3044 ATCAGTGTGGTGTGGTAAACGTAACATCTATAAGTCTATTTCTGCACTCTCTGTCAAA 2985  
QY 907 GCAGCAAAATCAGGCTAGTCTGCTGCAAGCTCTTAGAGTGGGAGTGGTTCCACACCC 966  
Db 2984 GCGGCTAAAGATGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2925  
QY 967 CAAGAGCTGAAGTGATGATAAGAACTTGGTTTGTGGCTACAGAGTCTCAACCTT 1026  
Db 2924 CAAATGCTGAAGTTAATGACAGCAAAATACGGTTTGTGTCATCAGAACGCTTGAACCCG 2865  
QY 1027 CAAAAAGCCAGAGTCTTCTTATGTTAGCCCTCACCACAACTAGTGTAGAGAGGCGATC 1086  
Db 2864 CAGAAAGCTGCTGTTCTTCTGCAATTTGGCTCTGACTCAGACTGCTGACACAGCAAAAATT 2805  
QY 1087 CAAAAGATCTTCTCCACCTATTATTCAGAA 1118  
Db 2804 CAGAAATGTTTCTTAAATATTAATTCCTGAA 2773

## RESULT 10

US-60-215-161-1294/c  
; Sequence 1294, Application US/60215161  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Huesing, Joseph E.  
; APPLICANT: Krasomil-Osterfeld, Karina C.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)A  
; CURRENT APPLICATION NUMBER: US/60/215,161  
; CURRENT FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 1294  
; LENGTH: 9082  
; TYPE: DNA  
; ORGANISM: Xenorhabdus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (6073)..(6924)  
US-60-215-161-1294

Query Match 26.7%; Score 302.8; DB 75; Length 9082;  
Best Local Similarity 58.1%; Pred. No. 3.4e-81;  
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;  
QY 127 CCCAAGTACTATCTAGCCACAGGAGGACCATCTGCTCGGGGAATCTAGCGTC 186  
Db 3755 CCAACATTACTGTTCTGGCAACGGGTGGCACATCGTGGTGGTGAATCTGCAACT 3696  
QY 187 AAGAGTAGTACTCTGCTGAGCAGTCCCGTTGATGAAGTCTTCTGAGCGGCTCCGCC 246  
Db 3695 AATCCAGCTACACTGCGGGTCAAGTCGGTGTGATTCATTCGTTGATTCGTTCTGCT 3636  
QY 247 ATCAAGCTAGTCTGCTGAGGAGGACAGATCTCAAGCTCTCTCGGCCCAAAAGATG 306

```

Db 3635 ATCAAAAACATTGCTAATCTGAAGGTGACGAGTGTGTTAGATCGGTCTTCAGGACATG 3576
QY 307 ACGGGTAAGGTGGCTTAACTAGCAAGCGGTGTAATGAGTCTTCGCGCCAAAAGAG 366
Db 3575 AACGATCAGGTCTGGCTGACTCTGGCGMAAAAATCAATG-----CAGATTGTGATAAA 3522
QY 367 ACCGAAGCGGTGATCATACCCATGAACATGACACCATGAAGAGAGCGGTCTTCCTC 426
Db 3521 ACTGACGGTGTGTCATACCCACCGGTACAGATACCATGAAGAAATCTGTTATTTCTC 3462
QY 427 AACCTCAGCGTGAAGAGCAAAACCTGCTGCTGCTAGCGCCATCGCTCCAGGCTCT 486
Db 3461 GATCTGACCACTCAGTGTCTAATACCTATCGTAACTGTTGGCGAAGTCCATCAACG 3402
QY 487 TCCATGAGTCTGATGCGCCCATGAAATCTCTAATAGCCCGTGAATGAGGATCAACAA 546
Db 3401 GCTCTGGGTGCTGATGCGCCCATGAACCTCTAATGCTGTTGTTAGCGTCAAGATAA 3342
QY 547 GCCTCTACTAACAAGAGGTGGTGTGATGATGAACGATGAGATTCAAGCGCCAGAGAA 606
Db 3341 GACTCTGCAACCGGTGGTGTGTTACTTGTCTGATGAATGATTCGTAATTCATGACGTG 3282
QY 607 GCGACCAAGCTCAACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAGAGTAAATC 666
Db 3281 ATCGGTAACTGAGCAACCTGAAGTTCAGGCTTCCAGGAGTGAAGCGGCCACAA 3222
QY 667 GGCACAGTCTAATGACCAAGTCTGATTTTCACTCAATCCGTTGACCTCACACCTT 726
Db 3221 GGCTTCATTAACCGGTAAAGTAACTACTACTCT---GCAGCACCAGTGAAGCAGAT 3165
QY 727 GCAAGTGTGTTGATATGACAAATCGAAGACTCCCGAGTCTGATATTTCTTTACGCT 786
Db 3164 AAGCTGTTTGTGATGTCAGAAACTGACTGAATGCGCAAAAGTAGTATTGTGTATAAC 3105
QY 787 CACCCCGATGATGATGTTTGTAGTCAATGACGCCCTTCAGGAGGAGCCAAAGGAATC 846
Db 3104 TACTCTAATGCTCCGATCTCCGTTTAAAGCAATTTGTAGAGATGATTATAAGGTATT 3045
QY 847 ATCCATGACGAGTGGCAATGGAACCTTTCCCTTTGACTCAAAATGCTCTTGAAAAA 906
Db 3044 ATCAGTGTGCTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2985
QY 907 GCGACCAAGCTCAACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAGGTAATC 666
Db 3281 ATCGGTAACTGAGCAACCTGAAGTTCAGGCTTCCAGGAGTGAAGCGGCCACAA 3222
QY 667 GGCACAGTCTAATGCGAAAGTCTGAGTATTTCCTCAATCCGTTCCAGCTCACACCTT 726
Db 3221 GGCTTCATTAACCGGTAAAGTAACTACTACTCT---GCAGCACCAGTGAAGCAGAT 3165
QY 1087 CABAAGATCTTCCACCTATTAAATCAAGAA 1118
Db 2804 CAGAAATGTTTCTAATATTAATCTCTGAA 2773

```

RESULT 11

```

US-60-215-161-1295/c
; Sequence 1295, Application US/60215161
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof

```

```

; FILE REFERENCE: 38-21(51847)A
; CURRENT APPLICATION NUMBER: US/60/215,161
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 1295
; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6985)..(9081)
; US-60-215-161-1295

```

```

Query Match      26.7%; Score 302.8; DB 75; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3.4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCCAAGTGACTATCTTAGCCACAGAGGACACATCGTGTTCGGGGGAATCTAGCGTC 186
Db 3755 CCAAAACATTACTGTTCTGGCAACGGGTGCAAAATCGTGTGGTGGTGAATCTGCAACT 3696
QY 187 AAGAGTAGTACTCTCTGAGGACGATCACCGTTGATAAGCTTCTTCAGCGCTCCCTGCC 246
Db 3695 AAATCCAGTACACTGCGGGTCAAGTCCGGTGTGATTCAATGCTGAATGCTGTCTCCCT 3636
QY 247 ATCAACGACCTAGCCACCATCAAGGTTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATG 306
Db 3635 ATCAAAACATTTGCTAATCTGAAGGTGAGCAGGTTGTTAGCATCGTTCCTCAGGACATG 3576
QY 307 ACGGTAAAGGTGGCTTAACTAGCCAAAGCGTGTCAATGAGTCTCTCGGCCCAAAAGAG 366
Db 3575 AACGATCAGGTCTGGCTGACTCTGCGCAAAAAAATCAATG-----CAGATTGTGATAA 3522
QY 367 ACCGAAGCGGTGATCATACCCATGCAACTGACACCATGCAAGAGAGCGGTCTTCTTCCTC 426
Db 3521 ACTGACGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3462
QY 427 AACCTCAGCGTGAAGAGCCAAACCTGCTGCTTGTAGGCGCCATCGTTCAGGCTCT 486
Db 3461 GATCTGACCACTCAGTGTCTAATAACCTATCGTAACTGTTGGCGCAATGCGTCCATCAACG 3402
QY 487 TCCATGAGTCTGATGCGCCCATGAAATCTCTAATAGCCCGTGAATGAGGATCAACAA 546
Db 3401 GCTCTGGGTGCTGATGCGCCCATTGAACCTCTAATGCTGTTGTTGTTGTTGTTGTTGTT 3342
QY 547 GCCTCTACTAACAAGAGGTGGTGTGATGTAAGCAAGATGAGATTCAAGGTTCAAGCGCCAGAGAA 606
Db 3341 GACTCTGCAACCGGTGGTGTGTTTACTTGTGTAAGATGATTCGTAATTCATGACGTGAC 3282
QY 607 GCGACCAAGCTCAACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAGGTAATC 666
Db 3281 ATCGGTAACTGAGCAACCTGAAGTTCAGGCTTCCAGGAGTGAAGCGGCCACAA 3222
QY 667 GGCACAGTCTAATGCGAAAGTCTGAGTATTTCCTCAATCCGTTCCAGCTCACACCTT 726
Db 3221 GGCTTCATTAACCGGTAAAGTAACTACTACTCT---GCAGCACCAGTGAAGCAGAT 3165
QY 727 GCAAGTGTGTTGATATGACAAATCGAAGACTCCCGAGTCTGATATTTCTTTACGCT 786
Db 3164 AAGCTGTTTGTGATGTCAGAAACTGACTGAATGCGCAAAAGTAGTATTGTGTATAAC 3105
QY 787 CACCCCGATGATGATGTTTGTAGTCAATGACGCCCTTCAGGAGGAGCCAAAGGAATC 846
Db 3104 TACTCTAATGCTCCGATCTCCGTTTAAAGCAATTTGTAGAGATGATTATAAGGTATT 3045
QY 847 ATCCATGACGAGTGGCAATGGAACCTTTCCCTTTGACTCAAAATGCTCTTGAAAAA 906
Db 3044 ATCAGTGTGCTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2985
QY 907 GCGACCAAGCTCAACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAGGTAATC 966
Db 2984 GCGGCTAAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2925
QY 967 CAAGAGGCTGAAGTGGATGAAGAAATCTGTTTGTGGTACAGAGAGTCTCAACCT 1026
Db 2924 CAAAATGCTGAAGTTAATGACAGCAAAATACGGTTTGTGTCACACCGTATTCCTTCGGTTTCTACT 2865
QY 1027 CAAAAGCCAGAGTCTTATGTTAGTCCCTCACCACCACTAGTGTAGAGAGCGGATC 1086
Db 2864 CAGAAAGCTCGTGTCTTCTGCAATTTGGCTCTGACTCAGACTGCTGACACAGCAAAAAAT 2805
QY 1087 CABAAGATCTTCCACCTATTAAATCAAGAA 1118
Db 2804 CAGAAATGTTTCTAATATTAATCTCTGAA 2773

```

QY 967 CAAGAGGCTGAAGTGGATGATAAGAACTTGGTTTGGCTGACAGAGAGTCTCAACCT 1026  
Db |||||  
Db 2924 CAAAATGCTGAAGTTAATGACAGCAAAATACGTTTGGTGCATCAGAACGCTGACCG 2865  
QY 1027 CAAAAGCCAGAGTCTTCTTATGTTAGCCCTCACCACAACTAGTGTAGAGAGCGATC 1086  
Db |||||  
Db 2864 CAGAAAGCTGCTGTTCTTCTGCAATTTGGCTCTGACTCAGACTGCTGACACAGCAAAAT 2805  
QY 1087 CAAAAGATCTTCTCCACCTATTATATCCAGAA 1118  
Db |||||  
Db 2804 CAGAAATGTTTCTAAATATTAATCTCGAA 2773

## RESULT 12

US-60-215-161-1296  
; Sequence 1296, Application US/60215161  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Huesing, Joseph E.  
; APPLICANT: Krasomil-Osterfeld, Katina C.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)A  
; CURRENT APPLICATION NUMBER: US/60/215,161  
; CURRENT FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 1296  
; LENGTH: 9082  
; TYPE: DNA  
; ORGANISM: Xenorhabdus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (8058)..(9080)  
US-60-215-161-1296

Query Match 26.7%; Score 302.8; DB 75; Length 9082;  
Best Local Similarity 58.1%; Pred. No. 3.4e-81;  
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCAAGTGTACTCTTCTAGCCAGAGGACCAATCGCTGGTTCGGGGAACTTAGCGTC 186  
Db |||||  
Db 5328 CCAAACTATCTGTCTGGCAACGGGTGGCAACATCGCTGGTGGTGAATCTGCACT 5387  
QY 187 AAGAGTGTACTCTCTGCGAGCAGTCACCGTTGATAGCTTCTTGACGCGTCCCTGCC 246  
Db |||||  
Db 5388 AATCCAGCTACTGCGGTCAAGTCGGTGTGATTCATTCGCTGAATGCTGTTCTGCT 5447  
QY 247 ATCAACGACTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATG 306  
Db |||||  
Db 5448 ATCAAAACACATCTTAATCTGAAAGGTGAGCAGGTGTTAGCATCGTTCTCAGGACATG 5507  
QY 307 ACGGTGAAGTGGCTTAACTAGCAACGGGTGTAATAGCTTCATAGCTTCCTCGCCCAAAAGAG 366  
Db |||||  
Db 5508 AAGATCAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5561  
QY 367 ACCGAAGCGGTGATCATCAACCATGGAATGACACCATGGAAGAGACCGCTTTCTCTCTC 426  
Db |||||  
Db 5562 ACTGACGGTTTCTCATCCACCGGTGACAGATACCATGGAAGAACTGCTTATTTCTCTC 5621  
QY 427 AACCTCAGGTGAAAGCCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486  
Db |||||  
Db 5622 GATCTGACCACTCAGTGTCTATAAACCTATCGTAATGGTTGGCGCAATGGCTCCATCAAG 5681  
QY 487 TCATCAGTGTGATGCGCCCATGATCTCTATAGCCCGTGAATGATGATGATGATGATGATGAT 546  
Db |||||  
Db 5682 GCTCTGGGTGCTGATGGCCCATGTAACCTCTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5741  
QY 547 GCCTCTACTAACAAGAGAGTGTGATTTGATGAACGATGAGATTCACGCGCCAGAGAA 606  
Db |||||

Db 5742 GACTCTGCAACCGTGGTGTCTTACTTGTATGAATGATCTGTAAATCATGGACGTGAC 5801  
QY 607 GCGACCAAGCTCAACACCGCAGTCAATGCAATTTGCTTGGCCCAACACAGGTAATAATC 666  
Db |||||  
Db 5802 ATCGGTAAACTGAGCAAACTGAAGTTTCAGGCTTTCAGGCAAGTGAACGAGGCCACAA 5861  
QY 667 GGCACAGTCTATTATGCGCAAGTGCAGTATTTTCACTCAATCCGTTGACACCTCACACCTT 726  
Db |||||  
Db 5862 GGCCTCATTCATAACCGTAAAGTAAACTACTACTCT---GCAGCACCGAGTGAAGACAGAT 5918  
QY 727 GCAAGTGAAGTTGATATTAGCAAAATCGAAGAACTCCCAAGAGTTCGATATTCTTTACGCT 786  
Db |||||  
Db 5919 AAAGCTGTTTTGATGTCAGCAAACTGACTGAACTGCCAAAGTAGTATTGTGTATAAC 5978  
QY 787 CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846  
Db |||||  
Db 5979 TACTCTAATGCTTCCGATCTGCCGTTAAAGCATTTGTAGAGAATGATTTATAAAGGTATT 6038  
QY 847 ATCCATGCAAGCATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAAT 906  
Db |||||  
Db 6039 ATCAGTGTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6098  
QY 907 GCAGCAAAATCAGGCTAGTCTGCTCGAAGCTCTAGAGTGGGCGAGTGGTTCACCAACC 966  
Db |||||  
Db 6099 GCGGCTAAAGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6158  
QY 967 CAAGAGGCTGAAGTGGATGAAGAACTTCCCTTTGACTCAAAATGCTCTTGAATAA 906  
Db |||||  
Db 6159 CAAAATGCTGAAGTTAATGACAGCAAAATACGGTTTGTGTCATCAGAGTCTCAACCT 1026  
QY 1027 CAAAAGCCAGAGTGTCTTCTTATGTTAGCCCTCACCACAACTAGTGTAGAGAGCGCATC 1086  
Db |||||  
Db 6219 CAGAAAGCTCGTGTCTTCTGCAATTTGGCTCTGACTCAGACTGCTGACACAGCAAAAT 6278  
QY 1087 CAAAAGATCTTCTCCACCTATTATATCCAGAA 1118  
Db |||||  
Db 6279 CAGAAATGTTTCTTAAATATTAATTTCTCTGAA 6310

## RESULT 13

US-60-215-161-1297  
; Sequence 1297, Application US/60215161  
; GENERAL INFORMATION:  
; APPLICANT: Corbin, David R.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Huesing, Joseph E.  
; APPLICANT: Krasomil-Osterfeld, Katina C.  
; APPLICANT: Malvar, Thomas M.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Spiridonov, Sergei  
; APPLICANT: Hinkle, Gregory J.  
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51847)A  
; CURRENT APPLICATION NUMBER: US/60/215,161  
; CURRENT FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 8409  
; SEQ ID NO 1297  
; LENGTH: 9082  
; TYPE: DNA  
; ORGANISM: Xenorhabdus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (6861)..(7958)  
US-60-215-161-1297

Query Match 26.7%; Score 302.8; DB 75; Length 9082;  
Best Local Similarity 58.1%; Pred. No. 3.4e-81;  
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;  
QY 127 CCCCAAGTGTACTCTTAGCCAGAGGACCAATCGCTGGTTCGGGGAACTTAGCGTC 186  
Db |||||  
Db 5328 CCAAACTATCTGTCTGGCAACGGGTGGCAACATCGCTGGTGGTGAATCTGCAACT 5387

QY	187	AGAGTAGTACTCTCTGCTGGAGCAGTACCGTTGATAGCTTCTTGAGCGCTCCCTGCC	246
Db	5388	AAATCCAGGTACACTCGGGTCAAGTGGTGTGATTCATTGCTGATGCTGTTCTGCTGCT	5447
QY	247	ATCAACGACCTAGCCACCATCAAGGGTGAAACAGATCTCAAGCATGCTGCCAAGAGATG	306
Db	5448	ATCAAAACATTGCTAATCTGAAAGGTGAGCAGGTTGTTAGCATCGGTTCTCAGGACATG	5507
QY	307	ACGGGTAAAGTGTGGCTTAACTAGCAAGCGTGTCAATGAGTCTCTCGCCCAAAAGAG	366
Db	5508	AACGATCAGGTCTGGCTGACTCTGGCGAAATAAATCAATG-----CAGATTGTGATAA	5561
QY	367	ACGAAACCGGTGATCATCAACCATGGAACCTGACACCATGGAAGAGACCGCTTCTCCTC	426
Db	5562	ACTGACCGTCTGCTCATCAACCATGGAACCTGACACCATGGAAGAGACCGCTTCTCCTC	5621
QY	427	AACCTACCGGTGAAAGCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	486
Db	5622	GATCTGACCACTCAGTGTCTATAACCTATCGTAATGTTGGCGCAATCGCTCCATCAACG	5681
QY	487	TCCATGAGTCTGATGCGCCCATGATCTCTATAACCGCTGATGATGATGATGATGATG	546
Db	5682	GCTCTGGTGTCTGATGCGCCCATGATCTCTATAACCGCTGATGATGATGATGATGATG	5741
QY	547	GCCTCTACTAACAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATG	606
Db	5742	GATCTGCAAAACCGTGTGTTTACTTGTCTATGATGATGATGATGATGATGATGATG	5801
QY	607	CGGACCAAGCTCAACCAACCGCAGTCAATGCTATGCTGCTGCTGCTGCTGCTGCTG	666
Db	5802	ATCGGTAAACTGAGCAAACTGAGTTCAGGCTTCCAGGCGATGATGATGATGATGATG	5861
QY	667	GGCAGCTCTATATGCAAGTCAAGTCAATGCTCAATCGCTTCCAGCTCACAACCTT	726
Db	5862	GGCTTCATTCATAACCGGTAAAGTAAACTACTACTCT---GCAGCACCACTGAAAGC	5918
QY	727	GCAAGTGAAGTTTGTATATTAGCAAAATCGAAGTCTCCCAAGTCTGATGATGATG	786
Db	5919	AAAGCTGTTTTTGTATGTCAGCAAACTGACTGAACTGCAAAAGTAGGATTGTTGAT	5978
QY	787	CACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	846
Db	5979	TACTCTAATGCTTCGATCTCGCGGTTAAAGCATTTGTAGAGATGATGATGATGATG	6038
QY	847	ATCCATGCAAGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCA	906
Db	6039	ATCAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	6098
QY	907	GCAGCAAAATCAGCGTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	966
Db	6099	CGGCTAAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	6158
QY	967	CAAGAGCTGAAGTGAATGATAAGAACTTGTGTTGTTGTTGTTGTTGTTGTTGTTG	1026
Db	6159	CAAAATGCTGAAGTGAATGATAAGAACTTGTGTTGTTGTTGTTGTTGTTGTTGTTG	6218
QY	1027	CAAAAGCCAGAGTCTCTTATGTTAGCCCTCACCACCAAACTAGTGTGATGAGGCG	1086
Db	6219	CAGAAAGCTGCTGTTCTCTGCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6278
QY	1087	CAAAAGATCTTCTCCACCTATTAATCCAGAA	1118
Db	6279	CAGGAATGTTTTTCTAAATATAATTCCTGAA	6310

Query Match 26.7%; Score 302.8; DB 75; Length 9082;

Best Local Similarity 58.1%; Pred. No. 3.4e-81;

Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY	127	CCCCAAGTACTATCTCTGCTGGAGCAGTACCGTTGATAGCTTCTTGCAAGCGCTCCCTGCC	186
Db	5328	CCAAACATTACTGTTCTTGGCAACCGGTGGCAATCGCTGCTGCTGCTGCTGCTGCTGCT	5387
QY	187	AAGAGTAGTACTCTCTGCTGGAGCAGTACCGTTGATAGCTTCTTGCAAGCGCTCCCTGCC	246
Db	5388	AAATCCAGGTACACTCGGGTCAAGTGGTGTGATTCATTGCTGAACTGCTGCTGCTGCT	5447
QY	247	ATCAACGACCTAGCCACCATCAAGGTGAAACAGATCTCAAGCATGCTGCCAAGAGATG	306
Db	5448	ATCAAAACATTGCTAATCTGAAAGGTGAGCAGGTTGTTAGCATCGGTTCTCAGGACATG	5507
QY	307	ACGGGTAAAGTGTGGCTTAAACTAGCCAAAGCGTGTCAATGAGTCTCTCCCAAAAGAG	366
Db	5508	AACGATCAGGTCTGGCTGACTCTGGCGAAATAAATCAATG-----CAGATTGTGATAA	5561
QY	367	ACGAAACCGGTGATCATCAACCATGGAACCTGACACCATGGAAGAGACCGCTTCTCCTC	426
Db	5562	ACTGACCGTCTGCTCATCAACCATGGAACCTGACACCATGGAAGAGACCGCTTCTCCTC	5621
QY	427	AACCTACCGGTGAAAGCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	486
Db	5622	GATCTGACCACTCAGTGTCTATAACCTATCGTAATGTTGGCGCAATCGCTCCATCAACG	5681
QY	487	TCCATGAGTCTGATGCGCCCATGATCTCTATAACCGCTGATGATGATGATGATGATG	546
Db	5682	GCTCTGGTGTCTGATGCGCCCATGATCTCTATAACCGCTGATGATGATGATGATGATG	5741
QY	547	GCCTCTACTAACAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATG	606
Db	5742	GATCTGCAAAACCGTGTGTTTACTTGTCTATGATGATGATGATGATGATGATGATG	5801
QY	607	CGGACCAAGCTCAACCAACCGCAGTCAATGCTATGCTGCTGCTGCTGCTGCTGCTG	666
Db	5802	ATCGGTAAACTGAGCAAACTGAGTTCAGGCTTCCAGGCGATGATGATGATGATGATG	5861
QY	667	GGCAGCTCTATATGCAAGTCAAGTCAATGCTCAATCGCTTCCAGCTCACAACCTT	726
Db	5862	GGCTTCATTCATAACCGGTAAAGTAAACTACTACTCT---GCAGCACCACTGAAAGC	5918
QY	727	GCAAGTGAAGTTTGTATATTAGCAAAATCGAAGTCTCCCAAGTCTGATGATGATG	786
Db	5919	AAAGCTGTTTTTGTATGTCAGCAAACTGACTGAACTGCAAAAGTAGGATTGTTGAT	5978
QY	787	CACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	846
Db	5979	TACTCTAATGCTTCGATCTCGCGGTTAAAGCATTTGTAGAGATGATGATGATGATG	6038
QY	847	ATCCATGCAAGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCA	906
Db	6039	ATCAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	6098
QY	907	GCAGCAAAATCAGCGTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	966
Db	6099	CGGCTAAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	6158
QY	967	CAAGAGCTGAAGTGAATGATAAGAACTTGTGTTGTTGTTGTTGTTGTTGTTGTTG	1026
Db	6159	CAAAATGCTGAAGTGAATGATAAGAACTTGTGTTGTTGTTGTTGTTGTTGTTGTTG	6218
QY	1027	CAAAAGCCAGAGTCTCTTATGTTAGCCCTCACCACCAAACTAGTGTGATGAGGCG	1086
Db	6219	CAGAAAGCTGCTGTTCTCTGCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6278
QY	1087	CAAAAGATCTTCTCCACCTATTAATCCAGAA	1118
Db	6279	CAGGAATGTTTTTCTAAATATAATTCCTGAA	6310

Query Match 26.7%; Score 302.8; DB 75; Length 9082;

Best Local Similarity 58.1%; Pred. No. 3.4e-81;

Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

RESULT 14

US-60-215-161-1298

Sequence 1298, Application US/60215161

GENERAL INFORMATION:

APPLICANT: Corbin, David R.

APPLICANT: Goldman, Barry S.

APPLICANT: Huesing, Joseph E.

APPLICANT: Krasomil-Osterfeld, Karina C.



Sequence 163, App  
Sequence 19647, A  
Sequence 24, Appl  
Sequence 16, Appl  
Sequence 31, Appl  
Sequence 30, Appl  
Sequence 7, Appl  
Sequence 1386, Ap  
Sequence 231, App  
Sequence 52, Appl  
Sequence 53, Appl  
Sequence 3945, Ap  
Sequence 19587, A  
Sequence 13856, A  
Sequence 34797, A  
Sequence 43649, A  
Sequence 87, Appl  
Sequence 4130, Ap  
Sequence 19352, A

US-09-655-378A-163  
US-60-487-610-19647  
US-60-494-568-24  
US-60-494-568-16  
US-09-655-378A-31  
US-09-655-378A-30  
US-10-426-124-7  
US-10-374-780A-1386  
US-10-612-783-231  
US-10-360-123A-52  
US-10-360-123A-53  
US-10-273-573-3945  
US-60-487-610-19587  
US-10-357-930-13656  
US-10-357-930-34797  
US-10-357-930-43649  
US-10-623-813-87  
US-10-603-113-4130  
US-60-487-610-19352

ALIGNMENTS

RESULT 1  
US-10-349-374-3  
; Sequence 3, Application US/10349374  
; GENERAL INFORMATION:  
; APPLICANT: Donald L Durden  
; TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES  
; ASPARAGINASE IN THE TREATMENT OF HUMAN  
; HEMATOLOGIC AND AUTOIMMUNE DISEASE

NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESS: Lyon & Lyon  
STREET: 633 West Fifth Street  
SUITE: 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq for Windows 2.0  
CURRENT APPLICATION DATA:  
FILING DATE: 21-Jan-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/094,435  
FILING DATE: Filed Herewith  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 60/049,085  
FILING DATE: June 9, 1997  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 234/274  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1133 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 27, 2003, 01:11:55 ; Search time 335.126 Seconds  
(without alignments)  
3071.976 Million cell updates/sec

Title: US-09-937-982-3  
Perfect score: 1133  
Sequence: 1 atgggcagcagccatcatca.....aagaaagggaattctttcac 1133

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 780815 seqs, 454324936 residues

Total number of hits satisfying chosen parameters: 1561630

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA New:  
1: /cgn2\_6/prodata/2/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/prodata/2/pna/US05\_NEW\_COMB.seq.\*  
3: /cgn2\_6/prodata/2/pna/US07\_NEW\_COMB.seq.\*  
4: /cgn2\_6/prodata/2/pna/US08\_NEW\_COMB.seq.\*  
5: /cgn2\_6/prodata/2/pna/US09\_NEW\_COMB.seq.\*  
6: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq.\*  
7: /cgn2\_6/prodata/2/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1129.8	99.7	1133	6	US-10-349-374-3
2	302.8	26.7	9082	5	US-09-897-516A-1293
3	302.8	26.7	9082	5	US-09-897-516A-1294
4	302.8	26.7	9082	5	US-09-897-516A-1295
5	302.8	26.7	9082	5	US-09-897-516A-1296
6	302.8	26.7	9082	5	US-09-897-516A-1297
7	302.8	26.7	9082	5	US-09-897-516A-1298
8	296.4	26.2	1830121	6	US-10-329-670-1
9	277.2	24.5	1041	6	US-10-603-114-1219
10	84.4	7.4	966	6	US-10-617-320-791
11	67.4	5.9	82	6	US-10-406-676-3
12	67.4	5.9	1067	1	PCT-US03-06661A-9
13	67.4	5.9	1074	1	PCT-US03-06661A-11
14	65.2	5.8	759	6	US-10-423-156-60
15	63.4	5.6	759	6	US-10-423-156-59
16	61.4	5.4	88	6	US-10-375-209A-40
17	45.	4.0	5347	1	PCT-US02-34768-1
18	45.	4.0	5347	1	PCT-US02-34769-17
19	44.4	3.9	1365	6	US-10-372-473-8
20	44.2	3.9	1386	6	US-10-281-092-3
21	43.8	3.9	2076	6	US-10-426-124-5
22	42.8	3.8	1371	6	US-10-281-092-7
23	38.4	3.4	740	6	US-10-311-473-9
24	38.4	3.4	1157	6	US-10-311-473-8
25	38.4	3.4	5309	6	US-10-622-220-14
26	38.4	3.4	5537	6	US-10-622-220-13

\_\_\_\_\_

547	QY	GCCTCTACTAAACAAAGAGGTGGTGTATGTGATGAACGATGAGATTCACGCCGCCAGAA	606
3341	Db	GACTCTGCAAAACCGTGGTGTGTTTTTACTTTGGCTATGAATGATTCGTAAATTCATGACGTGAC	3282
607	QY	GGACCAAGCTCAACACCAACCGCAGTCAATGCATTTGCTTCGCCCAACAACAGGTAAATC	666
3281	Db	ATCGGTAAACTGAGCACAACCTGAAGTTCAGGCTTTCCAGCAGTGAACGGCAGGCCACAA	3222
667	QY	GGCACAGTCTATTATATGGCAAAGTCGAGTATTTCACTCAATTCGGTTCGAAGCTCACACCTT	726
3221	Db	GGCTTCATTATCAACCGTAAAGTAAACTACTACTCT--GCAGCACACAGTGAAGCAGAT	3165
727	QY	GCAAGTGAGTTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTCGATATCTTTTACGCT	786
3164	Db	AAAGCTGTTTTTGATGTGACAAACTGACTGAACCTGCCAAANGTAGTATTTGTTATAAC	3105
787	QY	CACCCGATGATACTGATGTTTTTAGTCAATGCGAGCCCTTCAGGCAGGAGCAAGGAATC	846
3104	Db	TACTCTAATGCTTCCGATCTGCGGTTAAAGCAATTTGTAGAGAAATGATTATAAGGTATT	3045
847	QY	ATCCATGACGAGCATGGGCAATGGGAACCCCTTTCCCTTTGACTCAAAATGCTCTTGAAAA	906
3044	Db	ATCAGTGCCTGGTGTGGTAACGGTAAACATCTATAAGTCTATTCTGGACTCTCTGTACAA	2985
907	QY	GCAGCNAATCAGCGGTAGTCGTGCTGCGAAGCTCTAGATGGCAGTGGTTCCACCAC	966
2984	Db	CGCGCTAAAGATGGTGTGTGTGTGTTGCTTCCAACCGTATTCCTTTCCGGTTTTCAC	2925
967	QY	CAAGAGGCTCAAGTGGATGATAAGAACTTGGTTTTGTGGCTACAGAGAGTCTCAAGCT	1026
2924	Db	CAAAATGCTGAAGTTAATGACAGCAATACGGTTTTTGTGATCAGAACGCTTGAAACCG	2865
1027	QY	CAAAAACCCAGAGTGCCTTCTTATGTTAGCCCTCACCAAAACTAGTGTAGAGAGCGGATC	1086
2864	Db	CAGAAAGCTCGTGTCTTCTCGAATTGCTCTGACTCAGACTGCTGACACAGCAAAAATT	2805
1087	QY	CAAAAGATCTTCTCCACCTATTATTCACAGAA	1118
2804	Db	CAGGAAATGTTTTCTAAATATTTATTTCTCTGAA	2773

### RESULT 3

```

US-09-897-516A-1294/c
; Sequence 1294, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847) B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 1294
; LENGTH: 9082
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6073)..(6924)
; OTHER INFORMATION:
US-09-897-516A-1294

```

Qy	127	CCCCAAGTACTATCTTAGCCACAGAGGCCACCATCGCTGGTTTCGGGGGAATCTAGCGTC	186
Db	3755	CCAAACATATCTGTCTCGCAACGGGTGGCACAATCGTGGTGGTGTGATCTGCAACT	3696
Qy	187	AAGAGTAGTACTACTCTGTCTGGAGCAGTCACCGTTGATAAGCTTCTTGACGGCTCCCTGCC	246
Db	3695	AAATCCAGCTACACTGCGGTCACGTCGGTGTGATTCATTGCTGAATGCTGTTCCTGCT	3636
Qy	247	ATCAACGACTAGCCACCATCAAGGTGACAGATCTCAAGCATTTGGCTCCCAAGAGATG	306
Db	3635	ATCAAAAACATTCCTAACTGAAAGGTGACAGGTGTGAGCATCGTTCTCAGACATG	3576
Qy	307	ACGGGTAAAGGTGTGGCTTAAACTAGCCCAAGCGTGTCAATGAGCTCTCGCCCCAAAAAGAG	366
Db	3575	AACGATCAGGTCGTGCTGACTCTGGCGAAAAAATCAATG-----CAGATTGTGATAAA	3522
Qy	367	ACCGAAGCGGTGATCATCACCCATGGAACTGACACCATGAAGAGACCGCTTTCTTCCTC	426
Db	3521	ACTGACGGTTTTGTTCATCACCCACGGTACAGATACCATTGAAGAAATCTGCTTATTTCTC	3462
Qy	427	AACCTCACGGTGAAAGCCAAAAACCTGTCTGCCTTTGTAGCGCCATCGCTCCAGGCTCT	486
Db	3461	GATCTGACCACTCAGTGTCTATAAACCTATCGTAATGGTTGGCGCAATCGCTCCATCAACG	3402
Qy	487	TCCATGAGTGTGATGGCCCAATGATCTCTATAACGCCGTGATGTAGCATCAACAAA	546
Db	3401	GCTCTGGGTGCTGATGGCCCATTTGAACCTCTATAATGCTGTGTGTAGCGTCAGATAAA	3342
Qy	547	GCCTCTACTAAACAAAGGAGTGTGATGTGATGAAACGATGAGATTTCAGCGCCCGCAGAGAA	606
Db	3341	GACTCTGCAACCGTGGTGTTTTACTTGTCTATGAATGATCTGTAAATTCATGGACGTGAC	3282
Qy	607	GGGCAAGACTCAACACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAGGTAAAAATC	666
Db	3281	ATCGGTAAACTCAGCACAACTGAAGTTTCAGGCTTTCAGGCAGTGAACGCGAGGCCACAA	3222
Qy	667	GSCACAGTCTATTATGGCAAAGTCGAGTATTTCACTCAATCCGTTTCGACCTCACACCCCTT	726
Db	3221	GGCTTCNTTCATAACGGTAAAGTAACTACTACTCT---GCAGCACCACTGAAACGAGAT	3165
Qy	727	GCAAGTGATTTGATATTAGCAAAATCGAAGAACTCCCGCAGAGTCGATATTCTTTACGCT	786
Db	3164	AAAGCTGTTTTTGATCTGACAAACTGACTGAACTGCCAAAAGTAGGTATTGTGTATAAC	3105
Qy	787	CACCCCGATGATCTCATGTTTTAGTCAATCGACGCCCTTCAGGCAGGAGCCAAAGGAATC	846
Db	3104	TACTCTAATGCTTCGATCTGCGGTTAAGCATTTCTAGAGATGATTAATAAGGTAAT	3045
Qy	847	ATCCATGACGAGCATGGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTCTTCAAAA	906
Db	3044	ATCAGTGTCTGGTGTGGTAAACGGTAACATCTATAAGTCTATTCTGGACTCTCTGTACAA	2985
Qy	907	GCAGCAAAATCAGGCGTGTGCTGCTCGAAGCTCTAGAGTGGGAGTGGTTCCACCAAC	966
Db	2984	GGGCTAAGATGGTGTGTGTGTGTTTCGTTCCAAACCGTATTCTCTTCGGTTTCACTACT	2925
Qy	967	CAAGAGGCTGAAGTGGATGATAAGAAACTTGTGTTTTTGTGGCTACAGAGATCTCAACCT	1026
Db	2924	CAAAATGCTGAAGTTAATGACAGCAAAATACGGTTTTGTGTGATCAGAACGCCCTGAAACCG	2865
Qy	1027	CAAAAGCCAGAGTGTCTTATGTAGCCCTCACAAAACCTAGTGATAGAGGGCATC	1086
Db	2864	CAGAAAGCTCGTGTCTTTCTGCAATTCGGCTCTGACTCAGCTGTGACACAGCAAAAAT	2805
Qy	1087	CAAAAGATCTTCTCCACCTATTAAATCAAGAA	1118
Db	2804	CAGGAATGTTTTCTAAATATTAAATCTCTGAA	2773

#### RESULT 4

RESULT 4  
US-09-897-516A-1295/c  
; Sequence 1295, Application US/09897516A

Query Match 26.7%; Score 302.8; DB 5; Length 9082;  
Best Local Similarity 58.1%; Pred. No. 2.2e-91;  
Matches 576; Conservative 0; Mismatches 407; Indels 9;

## ; GENERAL INFORMATION:

; APPLICANT: Corbin, David R.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Huesing, Joseph E.  
 ; APPLICANT: Malvar, Thomas M.  
 ; APPLICANT: Krasomil-Osterfeld, Karina C.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Spiridonov, Sergei  
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
 ; FILE REFERENCE: 38-21(51847)B  
 ; CURRENT APPLICATION NUMBER: US/09/897,516A  
 ; CURRENT FILING DATE: 2001-06-29  
 ; PRIOR APPLICATION NUMBER: US 60/215,161  
 ; PRIOR FILING DATE: 2000-06-30  
 ; NUMBER OF SEQ ID NOS: 8415  
 ; SEQ ID NO 1295  
 ; LENGTH: 9082  
 ; TYPE: DNA  
 ; ORGANISM: Xenorhabdus sp.  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (6985)..(9081)  
 ; OTHER INFORMATION:  
 ; US-09-897-516A-1295

Query Match 26.7%; Score 302.8; DB 5; Length 9082;

Best Local Similarity 58.1%; Pred. No. 2.2e-91;

Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCNAGTACTATCTAGCCACAGGAGGACCAATCGTGTTCGGGGGAATCTAGCGTC 186  
 Db 3755 CCAACATTAATCTGTCTGGCAACGGGTGGCAATCGCTGGTGTGTGAATCTGCAACT 3696  
 QY 187 AAGAGTAGTACTCTGCTGGAGCAGTCACCGTTTCAATAGCTTCTTGCAGCGCTCCCTGCC 246  
 Db 3695 AATCCAGCTACACTGCGGTGCAAGTGGTGTGATTCATTCGTAATGCTGTTCTGCT 3636  
 QY 247 ATCAAGCAGCTAGCCACCATCAAGGGTGAACAGATCTCAAGCAITGGCTCCCAAGAGATG 306  
 Db 3635 ATCAAAACATTTGTAATCTGMAAGGTGAGCAGGTGTTGTAGCATCGGTTCCTCAGGACATG 3576  
 QY 307 ACGGTAAGGTGGCTTAACTAGCCAGGCTGCAATGAGTCTTCGTCGCCCAAAAGAG 366  
 Db 3575 AAGCATCAGCTGGCTGACTCTGGCGAAATAATCAATG-----CAGATTGTGATAAA 3522  
 QY 367 ACCGAAGCCGTGATCATCAACCCATGGAACTGACACCATGGAAGAGACCGCTTCTTCTCTC 426  
 Db 3521 ACTGACGGTTTGTGTCATCAACCCAGGTACAGATACCATGAGAACTGCTATTCTCTC 3462  
 QY 427 AACCTCAGGTGAAGCCAAACAAACCTGCTGTCGCTTGTAGCGCCATCGCTCCAGGCTCT 486  
 Db 3461 GATCTGACCACTCAGTGTGTCAAAACCTATCGTAATGTTGGCGCAATGCGTCCATCAAG 3402  
 QY 487 TCCATCAGTCTGATGCCCCCATGATCTCTATAACCCCGTGAATGTAGCGATCAACAAA 546  
 Db 3401 GCTCTGGTGTGATGGCCCATTTGAACCTCTATAATGCTGTGTGTGTAGCGTCAAGATAA 3342  
 QY 547 GCTCTACTAAACAAAGAGTGGTGATTTGATGAACGATGAGATTCACGGCGCCAGAGAA 606  
 Db 3341 GACTCTGCAAAACCGTGGTGTGTTTACTTGTGATGAATGATTCGTAATTCATGACGTGAC 3282  
 QY 607 GCGACCAAGCTCAACACCCAGCAGTCAATGCAATTTGCTTCGCCCAACACAGGTAATATC 666  
 Db 3281 ATCGGTAATCTGAGCAACATGAAGTTCAGGCTTTCCAGGCAGTGAACGAGGCCCAAA 3222  
 QY 667 GGCACAGTCTATTATGGCAAGTCGAGTATTTTCACTCAATCCGTTTCGACCTCACACCTT 726  
 Db 3221 GCGTTTCAATCAACCGTAAAGTAACTACTACTCT---CGACACCAAGTGAAGCAGAT 3165  
 QY 727 GCAAGTGAATTTGATATTAGCAAAATCGAAGAACTCCCCAGAGTCGATATTTCTTAGCGT 786  
 Db 3164 AAGAGCTGTTTTGATGTCAGCAAACTGACTGAATCGCAAAAGTAGGTATTGTGTATAAC 3105

QY 787 CACCCGATGATGATGATGTTTACTCAATGAGCCCTTCAGGAGGAGGCAAAAGGAATC 846  
 Db 3104 TACTCTAATGCTTCGATCTGCGGTTAAAGCAATTTGTAGAGAAATGATTATAAAGGTATT 3045  
 QY 847 ATCCATGACAGGATGGCAATGGAAACCCCTTTCCCTTTGACTCAAAATGCTCTTTGAAAAA 906  
 Db 3044 ATCAGTGTGCTGGTGTGTTGTAACGGTAACATCTATAAGTCTATTCTGGACTCTCTGTCA 2985  
 QY 907 GGAGCCAAATCAGGCGTAGTCTGCTCGAAGCTCTAGAGTGGGACAGTGTTCACCAACC 966  
 Db 2984 GCGGCTAAAGATGCTGTTGTTGTTGTTCCAAACCGTATTCCTTTCCGTTTCTCACTACT 2925  
 QY 967 CAAGAGCTGAAGTGAATGATGAAGAACTTGGTTTGTGCTACAGAGAGTCTCAACCT 1026  
 Db 2924 CAAATGCTGAAGTTAATGACGAAATACGGTTTGTGTCATCAGAACCGCTGAAACCG 2865  
 QY 1027 CAAAAAGCCAGAGTGTCTTATGTTAGCCCTCACCAAACTAGTATAGAGAGCGCATC 1086  
 Db 2864 CAGAAAGCTCGTGTCTTCTGCAATTTGGCTCTGACTCAGACTGCTGACACAGCAAAAT 2805  
 QY 1087 CAAAAGATCTTCTCACCTATTATTAATCAAGAA 1118  
 Db 2804 CAGGAATGTTTTCTAAATATTAAATTCCTGAA 2773

## RESULT 5

US-09-897-516A-1296

; Sequence 1296, Application US/09897516A

; GENERAL INFORMATION:

; APPLICANT: Corbin, David R.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Huesing, Joseph E.  
 ; APPLICANT: Malvar, Thomas M.  
 ; APPLICANT: Krasomil-Osterfeld, Karina C.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Spiridonov, Sergei  
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
 ; FILE REFERENCE: 38-21(51847)B  
 ; CURRENT APPLICATION NUMBER: US/09/897,516A  
 ; CURRENT FILING DATE: 2001-06-29  
 ; PRIOR APPLICATION NUMBER: US 60/215,161  
 ; PRIOR FILING DATE: 2000-06-30  
 ; NUMBER OF SEQ ID NOS: 8415  
 ; SEQ ID NO 1296  
 ; LENGTH: 9082  
 ; TYPE: DNA  
 ; ORGANISM: Xenorhabdus sp.  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (8058)..(9080)  
 ; OTHER INFORMATION:  
 ; US-09-897-516A-1296

Query Match

Best Local Similarity 58.1%; Pred. No. 2.2e-91;

Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCCAAGTACTATCTAGCCACAGGAGGACCAATCGTGTTCGGGGGAATCTAGCGTC 186  
 Db 5328 CCAACATTAATCTGTCTGGCAACGGGTGGCAATCGTGTGGTGTGTAATCTGCAACT 5387  
 QY 187 AAGAGTAGTACTCTGCTGGAGCAGTCACCGTTGATAGTCTTTGACCGCTCCCTGCC 246  
 Db 5388 AATCCAGCTACACTGCGGTGCAAGTGGTGTGATTCATTCGTAATGCTGTTCTGCT 5447  
 QY 247 ATCAAGCAGCTAGCCCAATCAAGGGTGAACAGATCTCAAGCAITGGCTCCCAAGAGATG 306  
 Db 5448 ATCAAAACATTTGTAATCTGAAAGGTGAGCAGGTGTTGTAGCATCGTCTCTCAGGACATG 5507  
 QY 307 ACGGTAAGGTGGCTTAACTAGCCAGCGTGTCAATGAGTCTCTGCCCAAAAGAG 366





```

; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c

```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:

```

; NAME/KEY: misc feature  
; LOCATION: (145171)..(145171)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (145942)..(145942)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (147197)..(147197)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (150841)..(150841)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (152500)..(152500)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (152530)..(152530)  
; OTHER INFORMATION: n equals a, t, g or c

Query Match 26.2%; Score 296.4; DB 6; Length 1830121;  
Best Local Similarity 57.1%; Pred. No. 4e-88;  
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;  
QY 126 ACCCCAGTACATCTCTAGCCACGAGGACCATCGCTGGTTCGGGGGATCTAGCGT 185  
DB 802722 ACCAAATATTACAACTTGGCAACGGGTGGTACCATTGCGAAGCGGGCAAGTTCGGT 802781  
QY 186 CAAGAGTAGTACTCTCTGAGCAGATCACCCTGATTAAGCTTCTGACGCGTCCCTGC 245  
DB 802782 AATTCTGCTGTAAGCTGACCAATTAAGTATTGATTAATTAAGCTGTACCA 802841  
QY 246 CATCAACGACTAGCCACCATCAAGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAT 305  
DB 802842 AATGAAAAATTTGCCAACATTAAGGTGAGCAAAATTTAAAAATAGTTTCAACAAGACAT 802901  
QY 306 GAGGGTAAGTGTGGCTTAACCTAGCCACGGTGTCAATGAGCTCCTCGCCCAAAAAGA 365  
DB 802902 GAATGACGAAGTCTGGCTAAACCTGGCAAAAGCCATCAATGCTC-----AATGTAAAG 802955  
QY 366 GACCGAAGCGTGTATCATCCCATGGAACCTGACACCATGGAAGACGCTTTCTCTCT 425  
DB 802956 TACTGATGATTTGTCTAATCCCATGTTACAGATACCAATGGAAGAGAGCGCTTATTTCTT 803015  
QY 426 CAACCTCAGCGTGAAAAAGCCAAAACCTGTCTGCTTTGAGCGCCATCGCTCCAGGCTC 485  
DB 803016 AGATTTAACCGTAAATGTGAAAAACCGTTGTTCTGTTGGGGCAATGCGTCTGCAAC 803075  
QY 486 TTCCATGAGTGTGATGGCCCATGAATCTCTATAACCGCTGAATGTAGCATCAACAA 545  
DB 803076 AGAAAAAGTGTGATGGCCCATTAATCTTTTACAATGCTGTCTGTTGCGAGCAGACAA 803135  
QY 546 AGCTCTTACTAAACAAAGAGTGGTATTGTGATGAACGATGAGATTTCACGCGCCAGAGA 605  
DB 803136 AATAATCAAGTGTGCTGTTTGTAGTGGCAATGAATGAATGATGATGCTGCTGCGA 803195  
QY 606 AGCGCAAGCTCAACACCCAGTCAATGCTTTGCTTTGCGCCCAACACAGGTAAAT 665  
DB 803196 TGTAACAAAAACAGTACGACGCGAGTCAAAACGTTCCATTACCAAAATATGTTCTCT 803255  
QY 666 CGGCACAGTCTATTATGGCAAGTCGAGTATTCTCACTCAATCCGTTGACCTCACACCT 725  
DB 803256 AGGCTATTTCATAACGCAAGTGGACTATGAACGTTTCCCGAAGAACATACCAT 803315  
QY 726 TGCAAGTGAAGTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTCGATATTCTTTACGC 785  
DB 803316 CAACATCTCGTTTAAACGTAGAAAAATTAGATAGCTACCCAAAGTGGGATTTATTATGC 803375  
QY 786 TCACCCGATGATGATGATGTTTGTAGTCAATGTCAGCCCTTCAGGCGAGGCAAGGAAT 845

DB 803376 TTATTCAAATGCACCTGTGCAACCATTAACGCAATTACTCAATGTGGTATCAAGGGAT 803435  
QY 846 CATCCATGCGAGGATGGCAATGGAACCCCTTTCCCTTTGACTCAAAATGCTCTTGAAAA 905  
DB 803436 TGATCTGCAGAGTTGGCAATGGAATGTTAATGCTGCACACTAGATCGGTAGAAAA 803495  
QY 906 AGCAGCCAAATCAAGGCGTAGTCTGCTCGCTCGAAGCTCTAGAGTGGCAGTGGTTCCACC 965  
DB 803496 AGCGCAAAAGATAGCGTCTGTTAGTCCGTTCTCTCGTGACCAACGGGTTATACAAC 803555  
QY 966 CCAAGAGCTGAAGTGAATGAATAAAGTCTGTTTGGTGGTACAGAGTCTCAACCC 1025  
DB 803556 TCGTGACGCTGAAGTGTATGATAGTAATATGCGTTTGTAGCATCAGGTACTTTAATCC 803615  
QY 1026 TCAAAAAGCCAGAGTCTCTTATGTTAGCCCTCAACAAACTAGTATAGTAGAGGCGAT 1085  
DB 803616 ACAGAAAGCCCGCTGCTCTTGCAATTAGCTTTAACTCAAACTAAAGATCCNAAAGTAT 803675  
QY 1086 CCAAAAGATCTTCTCCACCTATTAA 1110  
DB 803676 TCAACAATATTCGAAGACTTCTAA 803700

## RESULT 9

US-10-603-114-1219  
; Sequence 1219, Application US/10603114  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/10/603,114  
; CURRENT FILING DATE: 2003-06-24  
; PRIOR APPLICATION NUMBER: US/09/543,681  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 1219  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-10-603-114-1219

Query Match 24.5%; Score 277.2; DB 6; Length 1041;  
Best Local Similarity 56.5%; Pred. No. 3.7e-83;  
Matches 560; Conservative 0; Mismatches 423; Indels 9; Gaps 2;

QY 119 TGGCTAAACCCCAAGTGAATCTCTAGCCACGAGGAGCAACCATCGCTGGTTCGGGGGAAT 178  
DB 59 TTGCTTTACCTTAATGTGACATCTCGCAACTGGGGGAACCATCGCAGGGGTGGTGACT 118  
QY 179 CTAGCGTCAAGAGTAGTACTCTGCTGGAGCAGTCCCGTTGATAGCTTTTTCGACGCG 238  
DB 119 CTGCCACTTCTTCAAGCTACACCGCAGGAAAAATAGGCAATTGATACATTGATTAATGCCG 178  
QY 239 TCCTGCGCATCAACGACCTAGCCACCATCAAGGTGAACAGATCTCAAGCATTTGGTCCC 298  
DB 179 TACTGAGGCAAAAAAGTGGCTAACTTAACAGGCGAGCAAGTGGTTAATATTGGTTCCG 238  
QY 299 AAGAGATGAGCGGTAAAGGTGGCTTAACTAGCCAAGCGTGTCAATGAGCTCTCGCCC 358  
DB 239 AAGATATGAACGACCAAGTGGTTAAACTGGCGAATAAATAATTAACG-----GATT 292  
QY 359 AAAAAGAGACGGAAGCGTGTATCAACCATGGAACCTGACACCATGGAAGACCGCTT 418  
DB 293 GTGATAAAACAGACGCGCTTTGTGATAACCCATGTTGATACGATACGATGGAAGACGCGCT 352  
QY 419 TCTTCTCAACCTCACGCTGAAAGCCAAAACCTGTCTGCTCTGTAGCGCCATGGCTC 478  
DB 353 ATTTTCTTGATTTAAACACGATGTAATAAACCGGTGGTGTGATGGGGCAATGGCC 412

```

QY 479 CAGGCTCTCCATGAGTCTGATGGCCCCCATGAATCTCTATAACGCCGTGAATGAGCGA 538
Db 413 CAGCAACAGCATTAGGTGCTGATGGTCCATTAAACCTCTATAAATGCGGTGGTGGCAA 472
QY 539 TCACAAAGCCTCTACTAAACAAGGAGTGGTATGATGATGATGATGATGATGATGATG 598
Db 473 GTGATAAAGCATCTGAAATCGTGGTGTATTTAGTTACCATGAATTAATTCAGTGATG 532
QY 599 CCAGAGAAGCGACCAAGCTCAACACACCGCGAGTCAATGCAATTTGCTTCGCCCCACACAG 658
Db 533 GTAAAGATGTTGTCAAATGATACGACAGAGTGCAGCAATTCAGCGGATTAATGAG 592
QY 659 GTAAATCGGCACAGCTATTATATGCAAAAGTCGATTTTCACTCAATTCGCTTCGACCTC 718
Db 593 GTGCTCAAGCTTATGTTTCATGATGTTAAAGTACACTATTACACT--GCTCAACACCGC 649
QY 719 ACACCTTGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 778
Db 650 GCGCTGAGAAAGTTGCTTTGATGTCAGCAAAATTAACCGAATTAACCAAAAGTAGGTATT 709
QY 779 TTTACGCTCACCCGATGATCTGATGTTTGTAGTCAATGTCAGCCCTTCAGGCAGGAGCA 838
Db 710 TTTTAACTATGCTAATGCTGCTGATTTACCGGCTTAAGCGTTTATAGATAACCACTTA 769
QY 839 AAGGAATPCATCCATGAGGAGTGGCAATGGAAACCCCTTTCCCTTTTGACTCAAAATGCTC 898
Db 770 AAGGTATTGTTAGTGCAGGTGTCGCAATGTTAACTTATATTTGACATCTTAAATACTT 829
QY 899 TTGAAAGAGCAGCAAAATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 958
Db 830 TAGCTGATGTTGTAATAAAGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 889
QY 959 CCACCCCAAGAGCGTGAAGTGTGATGATAGAACTTTGTTGTTGTTGTTGTTGTTGTTG 1018
Db 890 TTACCAACAAAATGGTGAAGTAGATGATGCAAAATATGGCTTTATTCCTTCAGAACGTT 949
QY 1019 TCACCCCTCAAAAGCCAGAGTGTCTTATGTTAGCCCTTCACCAAACTAGTGATAGAG 1078
Db 950 TAAATCCACAGAAAGCAAGAGTGTATTAACAGTTATCTTTGACAGAAACACAGATCCG 1009
QY 1079 AGGGATCCAAAGATCTTCTCCACCTATTAA 1110
Db 1010 CCACGATCCAAAGAACTTTGAAAAATATTAA 1041

```

```

RESULT 10
US-10-617-320-791
; Sequence 791, Application US/10617320
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/617,320
; FILING DATE: 10-Jul-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/ 085131

```

```

; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 791:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...966
; SEQUENCE DESCRIPTION: SEQ ID NO: 791:
US-10-617-320-791
Query Match 7.4%; Score 84.4; DB 6; Length 966;
Best Local Similarity 51.9%; Pred.No. 7e-18;
Matches 190; Conservative 0; Mismatches 176; Indels 0; Gaps 0;
QY 369 CCAAGCCGTGATCATCACCCATGGAACCTGACACCATGGAAGAGACCGCTTTCTTCTCAA 428
Db 231 CGATGGAGTGGTATCACACACGGAACCGATCTTTAGAGAAACAGCCTATTTCTTGA 290
QY 429 CCTCAGGTGAAAAGCCAAAACCTGCTGCTTGTAGGCGCATGCGTCAGGCTCTTC 488
Db 291 TACCATGGAAGTTCCCATATGCTATCGTTCTAACAGGAGCCATGCGTAGCTCCAATGA 350
QY 489 CATGAGTGTGATGCCCCCATGATCTTATACGCGGTGAATGTAGCGATCAACAAGC 548
Db 351 GCTCGTAGTGTGTTTATTAATTTACCTAGTCTTTACGAGTGGCCAGTGATGACAG 410
QY 549 CTCTACTAAAGAGGAGTGGTGTGATGTAACGATGAGATTCACGCCGCCAGAGAGC 608
Db 411 GGCTGCTGACAAGGAGTTTGGTCTGTTATGAAGATGAATCCACGCTGCCAAGTATGT 470
QY 609 GACCAAGCTCAACACACCGCGAGTCAATGCAATTTGCTTCGCCCAACACAGGTAAATCGG 668
Db 471 CACCAAAACACATACGACTTAATGTAACACCTTCCAGACTTCCAAACACATGCGCCCTCGG 530
QY 669 CACAGTCTATTATGGCAAGTGGAGTATTTCACTCAATCGTTCGACCTCACACCTTGC 728
Db 531 TCTCATGTAAGAAACAGGAAATCTCTACTTCAAAAACAGCTGAACCTCGTTCGCTTTGA 590
QY 729 AAGTGA 734
Db 591 CCTTGA 596

```

```

RESULT 11
US-10-406-676-3
; Sequence 3, Application US/10406676
; GENERAL INFORMATION:
; APPLICANT: Structural Genomix, Inc.
; APPLICANT: Antonyamy, Stephen
; APPLICANT: Feil, Ingeborg
; APPLICANT: Buchanan, Sean
; APPLICANT: Post, Kai W.
; APPLICANT: Liu, Yi
; APPLICANT: Lorber, David
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF PAK4D KINASE
; TITLE OF INVENTION: PAK4KD

```

```

RESULT 13
PCT-US03-06661A-11
; Sequence 11, Application PC/TUS0306661A
; GENERAL INFORMATION:
; APPLICANT: Chou, Su-Yi
; TITLE OF INVENTION: Method of Producing Antigens
; FILE REFERENCE: SAMG/0002 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/06661A
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: 10/231,114
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,470
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,063
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,213
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,298
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Streptomyces cinnamomeus ATCC 11874
PCT-US03-06661A-11

```

```

Query Match          5.9%; Score 67.4; DB 1; Length 1074;
Best Local Similarity 98.6%; Pred. No. 4.2e-12;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCAGCAGCCATCATCATCATGACGCGGCTGTGTCGCGCGGCAGGCCAT 60
1 ATGGCAGCAGCCATCATCATCATGACGCGGCTGTGTCGCGCGGCAGGCCAT 60
61 ATGGCTAGC 69
61 ATGGCTAGC 69

RESULT 14
US-10-423-156-60
; Sequence 60, Application US/10423156
; GENERAL INFORMATION:
; APPLICANT: Lin, Hsin-Yu
; APPLICANT: Hwong, Ching-Long
; TITLE OF INVENTION: ANTIGENIC FRAGMENT OF HUMAN
; TITLE OF INVENTION: T-LYMPHOTROPIC VIRUS
; FILE REFERENCE: 05204-020001
; CURRENT APPLICATION NUMBER: US/10/423,156
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: TW 91135980
; PRIOR FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated oligonucleotide
US-10-423-156-60

```

Query Match 5.8%; Score 65.2; DB 6; Length 759;  
Best Local Similarity 95.7%; Pred. No. 2e-11;  
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
1 ATGGGCGAGCCATCATCATCATAGCAGCGCGTGTGCCCGGGCAGCCAT 60

[illegible]

Search completed: August 27, 2003, 11:03:32  
Job time : 341.126 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:28:33 ; Search time 4824.26 Seconds  
(without alignments)  
9607.814 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133  
Sequence: 1 atggcagcagccatcatca.....aagaaagggaatctcttcac 1133

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database :

GenEmbl :

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1129.8	99.7	1133	6	AR157944	AR157944 Sequence
2	1027.8	90.7	1035	1	WSANSAGEN	AX464427 Sequence
3	1027.8	90.7	2505	1	WSAJ2933	X83689 W.succinoge
4	1027.8	90.7	3606	1	WSAJ2933	X89215 W.succinoge
5	368.8	32.6	303450	1	AP005085	AP005085 Wolinella
6	323.2	28.5	303450	1	AP005085	AP005085 Vibrio pa
7	300	26.5	10893	1	AE015308	AE015308 Shigella
8	300	26.5	292906	1	AE016988	AE016988 Shigella
9	296.8	26.2	10749	1	AE005526	AE005526 Escherich
10	296.8	26.2	266658	1	AP002563	AP002563 Escherich
11	296.8	26.2	303121	1	AE016766	AE016766 Escherich
12	296.4	26.2	10173	1	U32758	U32758 Haemophilus
13	296.4	26.2	110000	6	AR274513_07	Continuation (8 of
14	296.4	26.2	110000	6	AR274513_08	Continuation (9 of
15	295.2	26.1	1530	1	ECOLASNI	Continuation (8 of
16	295.2	26.1	1643	1	ECOLASNI	M34277 E.coli L-as
17	295.2	26.1	10448	1	ECOLASNI	M34234 E.coli L-as
18	295.2	26.1	141744	1	ECU28377	AE000378 Escherich
19	294	25.9	1848	6	I73512	U28377 Escherichia
20	294	25.9	1848	12	U06943	I73512 Sequence 15
21	282.6	24.9	10769	1	AE013881	U06943 Synthetic c
22	282.6	24.9	210050	1	AE013881	AE013881 Versinia
23	276	24.4	22204	1	AE008842	AJ414148 Versinia
24	276	24.4	230050	1	AE008842	AE008842 Salmonell
25	276	24.4	301574	1	AE016844	AL627277 Salmonell
26	273.2	24.1	305961	1	AE016937	AE016844 Salmonell
27	270.4	23.9	10094	1	AE001498	AE016937 Bacteroid
28	259.2	22.9	1196	6	BD092628	AE001498 Helicobac
29	259.2	22.9	10923	1	AE000585	BD092628 Identific
30	230.4	20.3	304500	1	AP005953	AP000585 Helicobac
31	224.8	19.8	2450	1	ECSC	AP005953 Bradyrhiz
32	223.2	19.7	2837	6	A14577	X14777 Erwinia chr
33	223.2	19.7	2837	6	E01113	A14577 asparaginase
34	221	19.5	35000	1	AB000617	E01113 Genomic DNA
35	221	19.5	221160	1	BSUB0002	AB000617 Bacillus
36	215.2	19.0	1399	1	ECASN	Z99105 Bacillus su
37	210.8	18.6	1017	6	AR177619	X12746 Erwinia chr
38	209.2	18.5	1709	1	AF056495	AR177619 Sequence
39	207.6	18.3	314150	1	CJ11168X1	AF056495 Pseudomon
40	205.2	18.1	308015	1	AE016783	AL139074 Campyloba
41	203.2	17.9	110000	2	AC020884_1	AE016783 Pseudomon
42	200.4	17.7	9937	1	AE004563	Continuation (2 of
43	195.4	17.2	2353	1	PFL238710	AE004563 Pseudomon
44	178.8	15.8	20516	1	AE008866	AJ238710 Pseudomon
45	128.8	11.4	1337	8	SPLASPA	AE008866 Salmonell
						Y11944 S.pombe gen

ALIGNMENTS

RESULT 1  
AR157944  
LOCUS  
DEFINITION Sequence 3 from patent US 6251388.  
ACCESSION AR157944  
VERSION AR157944.1 GI:16219888  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1133)  
AUTHORS Durden,D.L.  
TITLE Utilization of Wolinella succinogenes asparaginase to treat  
diseases associated with asparagine dependence  
JOURNAL Patent: US 6251388-A 3 26-JUN-2001;



Db	421	TTCTCAACCTCAGCGTGAAAGCCAAAAACCTGCTCTTTGTAGGCGCCATGCGTCCA	480	gene	/clone lib="lambda EMBL3"
Qy	481	GGCTCTTCATGAGTCTGATGGCCCATGAATCTCTATAACGCGTGAATGATGCGGATC	540	CDS	20..1012
Db	481	GGCTCTTCATGAGTCTGATGGCCCATGAATCTCTATAACGCGTGAATGATGCGGATC	540		/gene="ansA"
Qy	541	AACAAAGCCTCTACTAACAAAGAGTGGTCAATGTGATGAACGATGAGATTACGCGCGC	600		20..1012
Db	541	AACAAAGCCTCTACTAACAAAGAGTGGTCAATGTGATGAACGATGAGATTACGCGCGC	600		/gene="ansA"
Qy	601	AGAGAAGCGACCAAGCTCAACACACGCGAGTCAATGCTTCGCGCCCAACACAGGT	660		/EC number="3.5.1.1"
Db	601	AGAGAAGCGACCAAGCTCAACACACGCGAGTCAATGCTTCGCGCCCAACACAGGT	660		/codon start=1
Qy	661	AAATCGGCAGCTCTATTATGCAAAAGTGCAGTATTCTCAATCCGTTGACCTCAC	720		/transl table=11
Db	661	AAATCGGCAGCTCTATTATGCAAAAGTGCAGTATTCTCAATCCGTTGACCTCAC	720		/product="L-asparaginase"
Qy	721	ACCTTCGAAAGTGGTTGATATTAGCAAAATCGAAGACTCCCGAGTTCGATATTCTT	780		/protein_id="CAA58658.1"
Db	721	ACCTTCGAAAGTGGTTGATATTAGCAAAATCGAAGACTCCCGAGTTCGATATTCTT	780		/db_xref="GI:758652"
Qy	781	TACCTCACCCGATGATCTGATGTTTGTAGTCAATGACGCGCTTCAGGCGAGGCAAA	840		/translation="MAKPQVITLITAGGTIAGSSSESVKSSYSAGAVTVDKLLAAVPAI
Db	781	TACCTCACCCGATGATCTGATGTTTGTAGTCAATGACGCGCTTCAGGCGAGGCAAA	840		NDLATIKGEQISSIGSEMTGKWLKLAQVNNVLAOKTEAVIITHTGDMETAEF
Qy	841	GGAATCATCCATCAGGCGATGGCAATGGGAACCCCTTCCCTTTGACTCAAAATGCTTT	900		LNLTVKQKPVVLGAMRPGSSMSADGPNLYNANVINKASIKNSKGVVIVMDEIHA
Db	841	GGAATCATCCATCAGGCGATGGCAATGGGAACCCCTTCCCTTTGACTCAAAATGCTTT	900		AREATKLTNTVANAPSPNTKIGTIVYKVEYPTQSRPHTLASEFDISKIELPRV
Qy	901	GAAGAAGCAGCAAAATCAGGCGTGTAGTCTGCTCGAAGCTCTAGAGTGGCGAGTTC	960		DILYHPDDTDLVNAALOAGAKGIHAGMGNPNPFLTQNALEKAAGKGVVAVRSR
Db	901	GAAGAAGCAGCAAAATCAGGCGTGTAGTCTGCTCGAAGCTCTAGAGTGGCGAGTTC	960		VSGSTTQBAEVDKLGTFVATESLNPKARVILLMLALTKTSDREAIKQIFSTY"
Qy	961	ACCACCCAGAGCGTGAAGTGGATGAAGAACTTGGTTTGGCTACAGAGTCTC	1020		
Db	961	ACCACCCAGAGCGTGAAGTGGATGAAGAACTTGGTTTGGCTACAGAGTCTC	1020		
Qy	1021	AACCTTCAAAAGCCAGAGTCTTCTTATGTTAGCCCTCACCAAACTAGTAGAGAG	1080		
Db	1021	AACCTTCAAAAGCCAGAGTCTTCTTATGTTAGCCCTCACCAAACTAGTAGAGAG	1080		
Qy	1081	GGATCCAAAGATCTTCTCCACCTATTAAATCCAAAGAAAGGAATCTCTTCC	1133		
Db	1081	GGATCCAAAGATCTTCTCCACCTATTAAATCCAAAGAAAGGAATCTCTTCC	1133		
RESULT 3					
WSANSAGEN					
LOCUS	WSANSAGEN		1035 bp	DNA	linear
DEFINITION	W.succinogenes ansA gene.				BCT 01-APR-1995
ACCESSION	X83689				
VERSION	X83689.1				
KEYWORDS	ansA gene; asparaginase.				
SOURCE	Wolinella succinogenes				
ORGANISM	Wolinella succinogenes				
	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;				
	Helicobacteraceae; Wolinella.				
REFERENCE	1				
AUTHORS	Derst, C. and Roehm, K.H.				
TITLE	Cloning and sequencing of L-asparaginase from Wolinella				
JOURNAL	succinogenes				
REFERENCE	2 (bases 1 to 1035)				
AUTHORS	Roehm, K.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-JAN-1995) K.H. Roehm, Physiologische Chemie der				
FEATURES	Philipps-Univ., Karl-von-Friesch-Str.1, D-35033 Marburg (Lahn), FRG				
	Location/Qualifiers				
	1..1035				
	/organism="Wolinella succinogenes"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:844"				

Qy	103	AGCGCTCTGTTTGTATGCTTAAACCCCAAGTACTATCTCTGCTGGAGCAGTCAACCTG	162		
Db	5	AGCGCTCTGTTTGTATGCTTAAACCCCAAGTACTATCTCTGCTGGAGCAGTCAACCTG	64		
Qy	163	GCTGTTGGGGGAATCTAGCGTCAAGAGTACTTCTCTGCTGGAGCAGTCAACCTG	222		
Db	65	GCTGTTGGGGGAATCTAGCGTCAAGAGTACTTCTCTGCTGGAGCAGTCAACCTG	124		
Qy	223	AAGCTTCTTTCGAGCGTCCCTCCATCAACGACGATAGCCATCAAGGTTGAAACAGATC	282		
Db	125	AAGCTTCTTTCGAGCGTCCCTCCATCAACGACGATAGCCATCAAGGTTGAAACAGATC	184		
Qy	283	TCAGCATTTGGCTTCCCAAGAGATGACGGGTAGGTTGGCTTAACTAGCCAAAGCGTGC	342		
Db	185	TCAGCATTTGGCTTCCCAAGAGATGACGGGTAGGTTGGCTTAACTAGCCAAAGCGTGC	244		
Qy	343	AATGAGTCTCTCGCCCAAAAGAGACCGAAGCGGTGATCATCAACCATGGAATGACAC	402		
Db	245	AATGAGTCTCTCGCCCAAAAGAGACCGAAGCGGTGATCATCAACCATGGAATGACAC	304		
Qy	403	ATGGAAGAGACCGCTTCTTCTCAACCTCAGCGTCAAGGTTGAAAGGCTGCTGCTGCT	462		
Db	305	ATGGAAGAGACCGCTTCTTCTCAACCTCAGCGTCAAGGTTGAAAGGCTGCTGCTGCT	364		
Qy	463	GTAGGCGCCATGCGTCCAGGCTCTTCCATGAGTGTGATGGCGCCCATGATCTCTATAAC	522		
Db	365	GTAGGCGCCATGCGTCCAGGCTCTTCCATGAGTGTGATGGCGCCCATGATCTCTATAAC	424		
Qy	523	GCGTGAATGTAGCGATCAACAAAGCTCTTCTAAACAAAGAGTGGTGGTGGTGGTGAAC	582		
Db	425	GCGTGAATGTAGCGATCAACAAAGCTCTTCTAAACAAAGAGTGGTGGTGGTGGTGAAC	484		
Qy	583	GATGAGATTCAACGCGCCAGAGAGCGACCAAGCTCAACACCCAGCGAGTCAATGCAATT	642		
Db	485	GATGAGATTCAACGCGCCAGAGAGCGACCAAGCTCAACACCCAGCGAGTCAATGCAATT	544		
Qy	643	GCTTCGCCCAACACAGGTAAATCGGACAGTCTTATATGGCAAGTGGATTTTCACT	702		
Db	545	GCTTCGCCCAACACAGGTAAATCGGACAGTCTTATATGGCAAGTGGATTTTCACT	604		
Qy	703	CAATCCGTTGACCTCACACCTTGCAGTCTTGTATATAGCAAAATCGAAGAACTC	762		
Db	605	CAATCCGTTGACCTCACACCTTGCAGTCTTGTATATAGCAAAATCGAAGAACTC	664		
Qy	763	CCGAGTGGATATTTCTTTAGCTCACCCCGATGATGATGATGATGATGATGATGATGATG	822		
Db	665	CCGAGTGGATATTTCTTTAGCTCACCCCGATGATGATGATGATGATGATGATGATGATG	724		

BASE COUNT 293 a 277 c 236 g 229 t

ORIGIN

Query Match 90.7%; Score 1027.8; DB 1; Length 1035;

Best Local Similarity 99.8%; Pred. No. 1e-284;

Matches 1029; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



QY	703	CAATCGGTTGCACTTCAACCCCTTCAAGTGAAGTTGATATAGCAAAATCGAAGAACTC	762
Db	1774	CAATCGGTTGCACTTCAACCCCTTCAAGTGAAGTTGATATAGCAAAATCGAAGAACTC	1833
QY	763	CCAGAGTCGATATCTTTAGCTCACCCTCGATGATCTGATGTTTATAGTCAATCGAGCC	822
Db	1834	CCAGAGTCGATATCTTTAGCTCACCCTCGATGATCTGATGTTTATAGTCAATCGAGCC	1893
QY	823	CTTACGAGGAGGCAAGGAATCATCTATGAGGATGGCAATGGAACCTTTCCTC	882
Db	1894	CTTACGAGGAGGCAAGGAATCATCTATGAGGATGGCAATGGAACCTTTCCTC	1953
QY	883	TTGACTCAAAATGCTCTTCAAAAGCAGCAATCAGGCGTAGTCTGCTCGAAGCTCT	942
Db	1954	TTGACTCAAAATGCTCTTCAAAAGCAGCAATCAGGCGTAGTCTGCTCGAAGCTCT	2013
QY	943	AGAGTGGCGAGTGGTTCCACCACCCCAAGAGGCTGAAGTGGATGAAGAACTTGGTTT	1002
Db	2014	AGAGTGGCGAGTGGTTCCACCACCCCAAGAGGCTGAAGTGGATGAAGAACTTGGTTT	2073
QY	1003	GTGGCTACAGAGTCTCAACCTTCAAAAGCCAGAGTGTCTTCTTATGTTAGCCCTCAC	1062
Db	2074	GTGGCTACAGAGTCTCAACCTTCAAAAGCCAGAGTGTCTTCTTATGTTAGCCCTCAC	2133
QY	1063	AAAACTAGTGATAGAGGCGGATCCAAAGATCTTCCACCTATTAAATCCAGAAAGG	1122
Db	2134	AAAACTAGTGATAGAGGCGGATCCAAAGATCTTCCACCTATTAAATCCAGAAAGG	2193
QY	1123	AATCTCTTAC	1133
Db	2194	AATCTCTTAC	2204
RESULT 5	WSAJ2933	3606 bp DNA linear	BCT 05-OCT-2000
LOCUS	Wolinella succinogenes aspA, dcuA genes and partial ansA gene.		
DEFINITION	Wolinella succinogenes aspA, dcuA genes and partial ansA gene.		
ACCESSION	AJ002933		
VERSION	AJ002933.1	GI:2644958	
KEYWORDS	ansA gene; aspA gene; aspartate ammonia-lyase; C4-dicarboxylate membrane transporter; dcuA gene; L-asparaginase.		
SOURCE	Wolinella succinogenes		
ORGANISM	Wolinella succinogenes		
REFERENCE	1		
AUTHORS	Ullmann, R., Gross, R., Simon, J., Unden, G. and Kroger, A.		
TITLE	Transport of C(4)-dicarboxylates in Wolinella succinogenes		
JOURNAL	J. Bacteriol. 182 (20), 5757-5764 (2000)		
MEDLINE	20461222		
PUBMED	11004174		
REFERENCE	2 (bases 1 to 3606)		
AUTHORS	Ullmann, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-NOV-1997) Ullmann R., Institut fuer Mikrobiologie, Johann-Wolfgang-Goethe Universitaet, Biozentrum N240, Marie-Curie-Strasse 9, 60439 Frankfurt, GERMANY		
COMMENT	Related sequence: X89215.		
FEATURES	Location/Qualifiers		
source	1..3606		
	/organism="Wolinella succinogenes"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:84"		
gene	400..1815		
	/gene="aspA"		
CDS	400..1815		
	/gene="aspA"		
	/codon_start=1		
	/transl_table=11		
	/product="aspartate ammonia-lyase"		
	/protein_id="CAA05764.1"		
	/db_xref="GI:2644959"		
db_xref="SPTREMBL:O34244"			
/translation="NASTRIEHLIDGLEVSNDCYVGVQTAQAKENFHSIGVLLSSMP			
TFIAGLAKVKKAAALANPELGLDEKIKNAICEACDLIIAGKYHQFVVDALOGGAGT			
STNNANEVIANVALELGHKKVEYKCHPNHNSLOSNDAYPTAFVDALEYKLIIE			
LTVSGILDESPAKAEEDFDIIMKRTLODAPVMTLQOEFTVAMTEIDIRVLE			
ANLVREMLGATAITGTGINSHPDIATLVEKLOQVTPFTASDLVEATQDTGAYV			
QISGVLRKSTVSKICNDLRLLSSGRAGLSEINLPMQPGSSIMPGKVNPIPRV			
NOVCVHIGADMTVTIASGGQLNPFVPIAFSSINMMRRAFETLAEKCVNGI			
TANADCRDYLKSGIQLTALNPYIGYENSTVAKALESGKSVHDIVLERKLLSKEE			
LEELIRPENWIKVVTLLTKK"			
1828..3129			
/gene="dcuA"			
1828..3129			
/gene="dcuA"			
/codon_start=1			
/transl_table=11			
/product="C4-dicarboxylate membrane transporter"			
/protein_id="CAA05765.1"			
/db_xref="GI:2644960"			
/db_xref="SWISS-PROT:O34245"			
/translation="MLVFEILVLAALFLVRLGGIAGVAGGLMIVLCGLGLKPG			
SIPDIVLIIIMSVIAAIAANQVAGGLDYLVLAEKLRSPQKHITFLAPVTVFMTLL			
AGTGHTAPSLPIAKVAKQGVKPSVLSIAVVSQIALITASPVSAVFMGALRP			
LGVYLOLLAICIPTTFIGMTAFICNLFNLDKDPVYQERVAKVAKLKGTOVY			
AKPKAKLSVLIFLVILAVFYATAISKWGLIQNPVGRDSAIMMFLMTATLIIAMF			
AKIDIVLNASTFKSGWTACICVLGAWLGDTEVSNHNDIKGFAGGILGHPMLIA			
ITLFFASMLLYSQATKALIPRAIALSVDPVTLIAFAVAASALFVLPTPTLLAAVQ			
MDDTGSTRIGKVFVNHPIVPEGLAIGISVALGFIVAPILL"			
3250..3606			
/gene="ansA"			
3250..3606			
/gene="ansA"			
/codon_start=1			
/transl_table=11			
/product="L-asparaginase"			
/protein_id="CAA05766.1"			
/db_xref="GI:2644961"			
/translation="NAKPOVTILATGTTAGSGSSVKSSVSAGAVTVDKLLAAVPAI			
NDAIIRKQIISIGSQEOMTGKWLAKRVNLLAQKETEAVIITHGTDTMEETAFF			
LNLTVKSRQKPVVLVGM"			
BASE COUNT	919 a	973 c	806 g 908 t
ORIGIN			
Query Match	32.6%	Score 368.8;	DB 1; Length 3606;
Best Local Similarity	99.5%;	Pred. No. 1.2e-94;	
Matches	370;	Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	103	AGCGCTCTGTTTGTGCTAAACCCCAAGTCACTATCTAGCCACAGGAGGCACCATC	162
Db	3235	AGCGCTCTGTTTGTGCTAAACCCCAAGTCACTATCTAGCCACAGGAGGCACCATC	3294
QY	163	GCTGTTTCGGGGGAATCTAGCGTCAAGAGTAGCTACTCTCTGGAGAGTCAACGTTGAT	222
Db	3295	GCTGTTTCGGGGGAATCTAGCGTCAAGAGTAGCTACTCTCTGGAGAGTCAACGTTGAT	3354
QY	223	AAGCTTCTTGACGCGTCCCTGCCATCAACGACTAGCCACCATCAAGGGTGAACAGATC	282
Db	3355	AAGCTTCTTGACGCGTCCCTGCCATCAACGACTAGCCACCATCAAGGGTGAACAGATC	3414
QY	283	TCAGCATTTGCTCCCAAGAGATCAAGGTAAGGTGTGGCTTAACTAGCCACGCTGTC	342
Db	3415	TCAGCATTTGCTCCCAAGAGATCAAGGTAAGGTGTGGCTTAACTAGCCACGCTGTC	3474
QY	343	AATGAGCTCTCGCCCAAAAAGAGACCGAAGCGGTGATCATCCCATGGAACCTGACACC	402
Db	3475	AATGAGCTCTCGCCCAAAAAGAGACCGAAGCGGTGATCATCCCATGGAACCTGACACC	3534
QY	403	ATGGAAGAGACCGCTTCTTCTCAACCTCACGGTGAAGAAAGCAAAAACCTGTCTGCCTT	462
Db	3535	ATGGAAGAGACCGCTTCTTCTCAACCTCACGGTGAAGAAAGCAAAAACCTGTCTGCCTT	3594
QY	463	GTAGGCGCCATG	474

Db	3595	GTAGGCCCATG	3606	gene	
RESULT 6	AP005085/c			CDS	
LOCUS	AP005085	303450 bp	DNA		linear BCT 05-MAR-2003
DEFINITION	Vibrio parahaemolyticus DNA, chromosome 2, complete sequence, 2/6.				
ACCESSION	AP005085	BA000032			
VERSION	AP005085.1	GI:28808465			
KEYWORDS					
SOURCE	Vibrio parahaemolyticus				
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;				
REFERENCE	1	Nasu, H., Iida, T., Sugahara, T., Yamauchi, Y., Park, K.S., Yokoyama, K., Makino, K., Shinagawa, H. and Honda, T.			
AUTHORS		A filamentous phage associated with recent pandemic Vibrio parahaemolyticus O3:K6 strains			
TITLE		J. Clin. Microbiol. 38 (6), 2156-2161 (2000)			
JOURNAL		20295086			
MEDLINE		10834969			
PUBMED					
REFERENCE	2	Makino, K., Oshima, K., Kurokawa, K., Kurokawa, K., Yokoyama, K., Uda, T., Tagomori, K., Iijima, Y., Najima, M., Nakano, M., Yamashita, A., Kubota, Y., Kimura, S., Yasunaga, T., Honda, T., Shinagawa, H., Hattori, M. and Iida, T.			
AUTHORS		Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism distinct from that of V. cholerae			
TITLE		Lancet 361 (9359), 743-749 (2003)			
JOURNAL		22508454			
MEDLINE		12620739			
PUBMED					
REFERENCE	3	(bases 1 to 303450)			
AUTHORS		Oshima, K., Kurokawa, K., Makino, K., Yokoyama, K., Yasunaga, T., Honda, T., Shinagawa, H., Hattori, M. and Iida, T.			
TITLE		Direct Submission			
JOURNAL		Submitted (09-APR-2002) Ken Kurokawa, Osaka University, Genome Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail: ken@gen-info.osaka-u.ac.jp, URL: http://www.gen-info.osaka-u.ac.jp, Fax: 81-6-6879-2047)			
COMMENT		genome project			
		This clone was isolated from a patient presenting with acute gastroenteritis.			
FEATURES		Location/Qualifiers			
source		1..303450			
		/organism="Vibrio parahaemolyticus"			
		/mol_type="genomic DNA"			
		/strain="O3:K6"			
		/sub_strain="RIMD 2210633"			
		/db_xref="taxon:670"			
		/chromosome="2"			
		81..1574			
		/gene="VPA0300"			
		81..1574			
		/gene="VPA0300"			
		/note="similar to EMB:CAC46187.1 (A1591787) percent identity 68 in 487 aa"			
		/codon_start=1			
		/evidence=not experimental			
		/transl_table=11			
		/product="hypothetical signal peptide protein"			
		/protein_id="BAC61643.1"			
		/db_xref="GI:28808466"			
		/translations="MKIKAVPKALLVTLASSLPISVAYAPNANESSILVQOQWEGN WPTQAEKLQVLYYORAIISAYQPMNVIGMRDGSFKGKGVNVLPIWKQMD RGLVPTPNADVVYNNYLDLKETGLVVAAPKVIQMTDFQRTIDVGLIGDRA RGLYLLLPNDVGPVKGYPFTSTNNVFLPFTVMKKGNGADPSDAVKIAEOTR IYPLWDEENIQPMKFPNAGORINMVPDTSYWEKMKQFDEPFISSAISPRLRGVL ASIGIVKGFPNFTVEEKKLKAUKTAPKMTAQKLIGKDKERDNLVYKQRYERAWA QYKLEHLPANPAKLFWAVTLINVDGTMPETSQLPSKNGFDNIENKTDGSDILVFGP QLPNGAPESNYIKTIIPGRDFLTALRYLGTGIEFFQTKWPKDDVLKIK"			
gene					
CDS					
		complement (1835..2314)			
		/gene="VPA0301"			
		complement (1835..2314)			
		/gene="VPA0301"			
		/note="similar to DBJ:BA043758.1 (AP003138) percent identity 57 in 158 aa"			
		/codon_start=1			
		/evidence=not experimental			
		/transl_table=11			
		/product="putative peptide methionine sulfoxide reductase"			
		/protein_id="BAC61644.1"			
		/db_xref="GI:28808467"			
		/translations="MMSEKIYPAGCCLMGVQFMRLHLPQVISTEAGRANGKTDNTQSEY DGVAECVTRTFDSQVSIETLMGFFIIPYSINKQDDVGEKRYTGVGRDTHKLE IAKSFIAITRADADKIAVELPLKNYVPSDEEHQDRLTRFPNDYCHIPLDLLHKYKNK" 2493..2882			
		/gene="VPA0302"			
		2493..2882			
		/gene="VPA0302"			
		/note="similar to GB:AAA64344.1 (L16865) percent identity 52 in 129 aa"			
		/codon_start=1			
		/evidence=not experimental			
		/transl_table=11			
		/product="hypothetical protein"			
		/protein_id="BAC61645.1"			
		/db_xref="GI:28808468"			
		/translations="MKIEHVAIWTERLEELKGFYEKYFNAVSNKYHNPKKHFSSYFL SPESGARLEMSMEGVTTGNSHMQVTGLAHFAFALGSEQAVDQITKLIVEDGYQRI DGPRTYGDGYESCVDLDPDGNRIELTV" 2974..3714			
		/gene="VPA0303"			
		2974..3714			
		/gene="VPA0303"			
		/note="similar to GB:AAL20622.1 (AE008775) percent identity 54 in 243 aa"			
		/codon_start=1			
		/evidence=not experimental			
		/transl_table=11			
		/product="putative regulatory protein"			
		/protein_id="BAC61646.1"			
		/db_xref="GI:28808469"			
		/translations="MNPONEILQIVNDRKRVQVTELSIIIGSVGTIRQDLNLEQQ GYLKRVHCAATALQSDIDTRLEVRFDIKQTLANKADLVAPNETVLIEGGSANALLA RTLAERGDVITITSAIHLIRNTSAILLGGVYHQGESLVGLTKLCIENTHFS TAFGLDGFQDGTGTRSDMMRADIAEAILAKKRNIVLTDSSKFGQIYPSIGKTNE ISVLLTDAAPKSLDLQLEKKGVEVLG" complement (3733..4089)			
		/gene="VPA0304"			
		complement (3733..4089)			
		/gene="VPA0304"			
		/note="similar to GB:AAF96321.1 (AE004376) percent identity 76 in 117 aa"			
		/codon_start=1			
		/evidence=not experimental			
		/transl_table=11			
		/product="conserved hypothetical protein"			
		/protein_id="BAC61647.1"			
		/db_xref="GI:28808470"			
		/translations="MHQHEKLVNVEFTGNIGATKAFHEHVGWQFVDYGDYAAFSG QSLDGFYSAEQVQTSNGALLVFYSSDIHATLEKXVAKFGQIIRPIPEFGGCRPH FTPEPSGNEFAVWSESH" 4319..5860			
		/gene="VPA0305"			
		4319..5860			
		/gene="VPA0305"			
		/note="similar to REF:NP_231225.1 (NC_002505) percent identity 83 in 510 aa"			
		/codon_start=1			
		/evidence=not experimental			
		/transl_table=11			
		/product="catalase"			
		/protein_id="BAC61648.1"			



## AUTHORS

Jin, Q., Yuan, Z.H., Xu, J.G., Wang, Y., Shen, Y., Lu, W.C., Wang, J.H.,  
Liu, H., Yang, J., Yang, F., Ou, D., Zhang, X.B., Zhang, J.Y., Yang, G.W.,  
Wu, H.T., Dong, J., Sun, L.L., Xue, Y., Zhao, A.L., Gao, Y.S., Zhu, J.P.,  
Kan, B., Chen, S.X., Yao, Z.J., He, B.K., Chen, R.S., Ma, D.L.,  
Qiang, B.Q., Wen, Y.M., Hou, Y.D. and Yu, J.

## TITLE

Genome sequence of *Shigella flexneri* 2a: insights into  
pathogenicity through comparison with genomes of *Escherichia coli*  
K12 and O157

## JOURNAL

Nucleic Acids Res. 30 (20), 4432-4441 (2002)

## PUBMED

12384590

## REFERENCE

2 (bases 1 to 10893)

## AUTHORS

Jin, Q., Shen, Y., Wang, J.H., Liu, H., Yang, J., Yang, F., Zhang, X.B.,  
Zhang, J.Y., Yang, G.W., Wu, H.T., Dong, J., Sun, L.L., Xue, Y.,  
Zhao, A.L., Gao, Y.S., Zhu, J.P., Chen, S.X., Yao, Z.J., Wang, Y.,  
Lu, W.C., Qiang, B.Q., Wen, Y.M. and Hou, Y.D.

## TITLE

Direct Submission

## JOURNAL

Submitted (08-MAY-2001) Microbial Genome Center of Chinese Ministry  
of Public Health, 100 YingXin Jie, XuanWu Qu, Beijing 100052, P.R.  
China

## FEATURES

Location/Qualifiers

1..10893  
/organism="Shigella flexneri 2a str. 301"  
/mol\_type="genomic DNA"  
/strain="301"  
/serotype="2a"  
/db\_xref="taxon:198214"  
188...541  
/locus\_tag="SF2947"  
188...541  
/locus\_tag="SF2947"  
/notes="Residues 1 to 114 of 117 are 80 pct identical to  
residues 1 to 114 of a 216 aa protein from *Klebsiella*  
*pneumoniae* gb: AAF06334.1"  
/codon\_start=1  
/transl\_table=11  
/product="orf, partial conserved hypothetical protein"  
/protein\_id="AAN44428.1"  
/db\_xref="GI:24053359"

## gene

## CDS

/translation="MKRSFTSLVCASLLALVSGVNAEKVALKLAHNERSHVHVH  
QSFELAKEVKOLSKGNVIRYPSSQGNARETMELLQNGALDWTGKSASDLESFDN  
IYAIYNLPILSCFW"  
558...1151  
/locus\_tag="SF2948"  
558...1151  
/locus\_tag="SF2948"  
/notes="Residues 2 to 197 of 197 are 48 pct identical to  
residues 145 to 341 of a 341 aa protein from *Bacillus*  
*halodurans* ref: NP\_241567.1"  
/codon\_start=1  
/transl\_table=11  
/product="orf, partial conserved hypothetical protein"  
/protein\_id="AAN44429.1"  
/db\_xref="GI:24053360"

## gene

## CDS

/translation="MDSTKDKGFALSAVAGTTSFYAKPIPKPDLKGLKIRVOPS  
PTTKMELMGSGPTPISEGVYTMQGVVDGAEENVPSWQTRHIEIAKVFSESDH  
ASIPDELAIVSTKWNKLTPEQQILETAARKSEAYOQKLWEKIDADTRAQAKAMGGEI  
KYDKAPRAVQPLFDFFDKKQKQALEKFDNAQ"  
1352...1699  
/gene="dctQ"  
1352...1699  
/locus\_tag="SF2949"  
1352...1699  
/gene="dctQ"  
/locus\_tag="SF2949"  
/notes="Residues 1 to 108 of 115 are 75 pct identical to  
residues 22 to 129 of a 136 aa protein from *Klebsiella*  
*pneumoniae* gb: AAF01413.1"  
/codon\_start=1  
/transl\_table=11  
/product="putative small integral C4-dicarboxylate  
membrane transport protein DctQ"  
/protein\_id="AAN44430.1"  
/db\_xref="GI:24053361"  
/translation="MMIIVGVLGAAYTVGAQRHSLDLPALANKRQLLSIVINVL

## gene

## CDS

/translation="MLMTGNAGWGEPPPTENILKQFKQYHGLKLDALSLKNLDAK  
GNQATVSGSDSDLYTVWQADLYELLEQWTWKDKPKVFSAMLSKGTSPASWS  
VNFYSFOAAASDRGRVDDIKTNKYLVINSEDFNYRFSQLESALNTQKNSIPALEKE  
ILFRAGSVIVTGGGLKLDKTLATSQVSAAMOIPMGYVYVLIPLSGVMMFALCFINQ  
SIQQLKQPAQAS"  
complement (1722..2117)  
/gene="insB"  
/locus\_tag="SF2950"  
complement (1722..2117)  
/gene="insB"  
/locus\_tag="SF2950"  
/notes="Residues 1 to 131 of 131 are 100 pct identical to  
residues 37 to 167 of a 167 aa protein from *Shigella*  
*flexneri* ref: NP\_085393.1"  
/codon\_start=1  
/transl\_table=11  
/product="orf, conserved hypothetical protein"  
/protein\_id="AAN44431.1"  
/db\_xref="GI:24053362"

## gene

## CDS

/translation="MIVCAEMDEHWGVGAKSRQRLFYAYDIRRRTVVAHVFGERTL  
ATUERLLSLSAFEVVMWTDGWPPLYESRLKGLHVISKRYTORIERHNLRLQHLAR  
LGRKSLFSKSVELDKVICHYLNKHQ"  
complement (2144..2419)  
/gene="insA"  
/locus\_tag="SF2951"  
complement (2144..2419)  
/gene="insA"  
/locus\_tag="SF2951"  
/notes="Residues 1 to 91 of 91 are 98 pct identical to  
residues 1 to 91 of a 91 aa protein from *Shigella flexneri*  
ref: NP\_085394.1"  
/codon\_start=1  
/transl\_table=11  
/product="orf, conserved hypothetical protein"  
/protein\_id="AAN44432.1"  
/db\_xref="GI:24053363"

/translation="MASISIRCPSCSATEGVVRNGKSTAGHORVLCSHCRKTWQLOFT  
YTAQPGTHQKIIDMANGVGCASRIMVGLNVLRLHKNRGR"  
2578...3777  
/locus\_tag="SF2952"  
2578...3777  
/locus\_tag="SF2952"  
/notes="Residues 1 to 395 of 399 are 61 pct identical to  
residues 38 to 431 of a 435 aa protein from *Salmonella*  
*typhimurium* LT2 ref: NP\_462086.1"  
/codon\_start=1  
/transl\_table=11  
/product="putative integral membrane protein, possible  
transporter"  
/protein\_id="AAN44433.1"  
/db\_xref="GI:24053364"

/translation="MLMFPWDIAAITVSORLANGLONFALLAIPFFIFAGTLMNSGG  
IAIRLINLAQVMVGRVPGSLGHVNVLANMFGSISGSAAVAAAAGVGLNPIQREGY  
DPAFSTAVNVSSCITGLIIPSNVLIIVPSLTAGSVASVAFMAGYLPGLMGLAMIV  
CGIIAKRGVPLSERATPAQAKAFDALPSLILVIVMGGILGFIPTATEASAIIV  
YRILSVLIYREVKNRHLPLILDESVVTTISVILLGFGSVGMAMTADIPYMISA  
LMGISNPLIILILINIVLIVGFMIDMTFAVIFTFPIAQLGMDPVPFGIMV  
AUNICILLTPPVGSALFVSGISGVKLIKELLLFYAALLVALLMVTITPQISLFI  
PQLLGLM"  
complement (3878..4804)  
/gene="yggM"  
/locus\_tag="SF2953"  
complement (3878..4804)  
/gene="yggM"  
/locus\_tag="SF2953"  
/notes="Residues 1 to 308 of 308 are 99 pct identical to  
residues 13 to 320 of a 335 aa protein from *Escherichia*  
*coli* K12 ref: NP\_417431.1"  
/codon\_start=1  
/transl\_table=11  
/product="putative alpha helix chain"  
/protein\_id="AAN44434.1"  
/db\_xref="GI:24053365"

/translation="MLMTGNAGWGEPPPTENILKQFKQYHGLKLDALSLKNLDAK  
GNQATVSGSDSDLYTVWQADLYELLEQWTWKDKPKVFSAMLSKGTSPASWS  
VNFYSFOAAASDRGRVDDIKTNKYLVINSEDFNYRFSQLESALNTQKNSIPALEKE



[illegible]

CDS	/gene="ptsp" /locus tag="S3037" complement(7274..9520) /gene="ptsp" /locus tag="S3037" /function="enzyme; Transport of small molecules: Other" /notes="residues 1 to 736 of 748 are 99.59 pct identical to residues 1 to 736 of 748 from Escherichia coli K-12; B2829" /codon_start=1 /transl_table=11 /product="PTS system, enzyme I, transcriptional regulator (with NTP and NTP proteins)" /protein_id="AAP18151.1" /db_xref="GI:30042427" /translations="MLTRURREIVKVASAPRLNEALNLIIVDICLAMDTEVCVSLAD HRCRYLMAITGLKPRGRTVTLAFDEGIVGLVRLAEPINLAQKRPSPKIPSV KEERFRAFLVPIIQRRLGLVQVQRELQRYDESEFLVTLTQMAAILQSQSLT ALFGVQROTRIRALPMQGVIAEAWQDQALPLMEQVYQASTLPALERLITGALAE AANEFERYSRFAAGKETAALFDLYSHLLSDRLRLRELFAEDVKGSAEWAQKTVI EKTAEPQALSDNYLKERAGDLRALGORLLPHLDDANOGPNAMPFERFILVADELSATT LAELPQRLVGVVRGAANSHAAIIVRALGIPYVMGADIQPSVLHRRLLTVGVYRGE LLVDPEVLQYORLSEIEUSLEIAEDDVNLPAPQKSGERIKVNLNAGLSPEHEK LGSCIDIGILYRTEIIPMLQSGFSEEEVAQYQGMFNDKRPVTLRLVDGAKQL PYMPSIENPGLRGIRITLDQPEIFLIQVRLAMRLANATNLNILLPVLSTLSDVD EARLIERACREVEEMGYIPIKPRIGIOMLEVSPVMFLPLAKRDFIVSVGNIDTQ YILAVDRNRTVANIYDSLHPAMLRLALAMTAREAEIHGIDRLCGEMAGDPMCVAILI
Query Match	26.5%; Score 300; DB 1; Length 292906;
Best Local	Similarity 57.1%; Pred. No. 1.6e-74;
Matches	568; Conservative 0; Mismatches 420; Indels 6; Gaps 1;
QY	119 TGGCTAAACCCCAAGTGACTCTCTAGCACACAGAGGACCATCGCTGGTTTCGGGGGAAT 178
DB	120985 TGGCATACCAATATACCAATTTAGCAACCGCGGGACCATTTGCCGGTGGTGACT 120926
QY	179 CTAGCGTCAAGAGTAGTACTCTGCTGAGCAGTCAACGTTGATAGCTTCTTCAGCGG 238
DB	120925 CGCAACCAAAATCTAACTACACAGCGGTAAAGTTGGCGTAGAATAATCTGGTTAATCGG 120866
QY	239 TCCCTGCCATCAACGACCTAGCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCC 298
DB	120865 TGCCGCACTAAAGACATTTGGACCTTAAGCGCAGCAGTAGTGAATCGCTCCC 120806
QY	299 AAGAGATGACGGGTAAAGTGTGGCTTAACTAGCACAACGGTGTCAATGAGTCTCTCGCC 358
DB	120805 AGGACATGAACGATAATGTCTGGCTGACACTGGCGCAAAAAAATTA-----CGCCGACT 120752
QY	359 AAAAGAGACCAAGCCGTGATCATCCCACTGGAACCTGACACCATGGAAGACCGCTT 418
DB	120751 GCGATAAACCAGCGCTTCTGTCATTAACCGTACCGACGATGGAAGAACTGCTT 120692
QY	419 TCTTCTCTCAACCTCAGGTGAAAAGCAAAAAACCTGTCTGCGTTGTAGGGCGCATGGCTC 478
DB	120691 ACTTCTCGACCTGACGGTGAATTCGCAAAACCGGTGGTGTGTCGCGCAATCGGCC 120632
QY	479 CAGGCTCTTCATGAGTGTGATGCCCATCAATCTCTATAAACCGCTGAATGATGACGA 538
DB	120631 CGTCCACGTCATGAGCGAGACGGTCCATTCAACCTGTATACCGGTAGTAGCCGAG 120572
QY	539 TCAACAAGCCTCTACTAAACAAGGAGTGGTGAATTTGTGATGAACGATGAGATCAAGCGG 598
DB	120571 CTGATAAAGCCTCCGCTAATCGTGGGTGCTGTGTGTGTATGAACACCGTACTGGACG 120512
QY	599 CCAGAGAGGACCAAGCTCAACACACCGCAGTCAATGCAATTTGCTTCCGCCAACACAG 658
DB	120511 GTCCGATGTGTACCAAAAAACCAACACCGCAGTAGAGCACCCTTCAAGTCTGTTAATACG 120452
QY	659 GTAAATCGSCACAGTCTATTATGGCAAAAGTCAGATTAATCAATCCGTTTCGACCTTC 718
DB	120451 GTCTCTGGGATACATTCACACGGTAAGATTGATTACCAACGTAACCCCGCAGGTAAAGC 120392
QY	719 ACACCCCTTGAAGTGGTGTGATTAATGAGCAAAATCGAAGAACTCCCAAGATCGATATTC 778
QY	

Db	120391	ACACCGGATACGCCATTGATGTCTCTAAGCTGAATGAGTCCGCAAGTTCGGCATCG	120333	
QY	779	TTTAGCTCACCAGATGATGATGTTTGTAGTCAATGACGAGCTTCAGCAGAGGCA	838	
Db	120331	TTTATAACTACGCTAAGCATCCGATCTTCGGGCTAAGCACTGTTAGTATCGGGCTATG	120272	
QY	839	AAGGAATCATCATCGACGAGCATGGCAATGGAAACCTTTCCCTTTGACTCAAAATGCTC	898	
Db	120271	ATGGCATGTTTAGCGCTGGTGTGGTAATGTAATGTAATAAATCCGTGTTCGACACCC	120212	
QY	899	TGAAAAAGCAGCAAAATCAGCGTAGTCTCGCTCGAAGCTCTAGAGTGGCGAGTGGTT	958	
Db	120211	TGGCAACCGCGCGAAAAACCGCACTGCAGTAGTGGCTCTTCCCGGTACCGAGCGGTG	120152	
QY	959	CCACCACCAAGAGAGCTGAAAGTGGATGATAAGAACTTTGGTTTGTGGCTACAGAGATC	1018	
Db	120151	CTACCACTCAGATGCTGAAGTGGATGATGCGAATACGGTCTGCTCGCTCTGCGACGC	120092	
QY	1019	TCAACCTCTAAAAGCCAGAGTCTTATGTTAGCCCTCACCAAAACCTAGTGATAGAG	1078	
Db	120091	TGAACCCGCAAAAGCGCGCTCTGCTGCGAGTGGCTCTGACGCAAAACCAAGATCGC	120032	
QY	1079	AGCGCATCAAAAGATCTTCTCCACTTAATC 1112		
Db	120031	ACAGATCAGCAGATTTCAATCAGTACTAATC 119998		
RESULT 9	AE005526/c	10749 bp	DNA linear	BCT 21-MAR-2001
LOCUS	Escherichia coli O157:H7 EDL933 genome, contig 3 of 3, section 145			
DEFINITION	of 290.			
ACCESSION	AE005526	AE005174		
VERSION	AE005526.1	GI:12517505		
KEYWORDS	Escherichia coli O157:H7 EDL933			
SOURCE	Escherichia coli O157:H7 EDL933			
ORGANISM	Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.			
REFERENCE	1 (bases 1 to 10749)			
AUTHORS	Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamousis, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.			
TITLE	Genome sequence of enterohaemorrhagic Escherichia coli O157:H7			
JOURNAL	Nature 409 (6819), 529-533 (2001)			
MEDLINE	21074935			
PUBMED	11206551			
REFERENCE	2 (bases 1 to 10749)			
AUTHORS	Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grobeck, E.J., Davis, N.W., Lim, A., Dimalanta, E., Potamousis, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA			
FEATURES	Location/Qualifiers			
source	1. 10749			
	/organism="Escherichia coli O157:H7 EDL933"			
	/mol_type="genomic DNA"			
	/strain="EDL933"			
	/serotype="O157:H7"			
	/db_xref="taxon:155864"			
	/note="enterohaemorrhagic"			

```

/gene="ansB"
/function="enzyme; Degradation of small molecules: Amino acids"
/note="Residues 1 to 348 of 348 are 98.85 pct identical to residues 1 to 348 of 348 from Escherichia coli K-12 Strain MG1655: B2957"
/codon_start=1
/transl_table=11
/product="periplasmic L-asparaginase II"
/protein_id="AAG58088.1"
/db_xref="GI:12517506"
/translation="MEFFPKTALAALVMGFSGAALALPNITILATGTTAGGDSATK SYTAGKGVENLVNAPQLKDIAVNGEQQVNIQSDMDDVWLTLAKKINTDCDKT DGFVTHGTDTEETAYFLDLTKCDKPPVMVGAMRPTSTMSADGPNLYXNAVTPAAR KASANGVLVNDVLDQRDVTNTKTNDVATFKSVNGPLGYHNGKIDVQRTPAAR HTSDTPFDVSKLNEPLKPGVIVYVNASDLPKALVDAGYDGVISAGVNGNLYKSVF DTLAAKNGTAVPSSRPVPTGATTQDAEVDADKYGFASGTLPNPKARVLLQLALTQ TKDPOIQIIFNOY"
/complement(13111..2030)
/gene="yggN"
/note="synonym: Z4303"
/complement(13111..2030)
/gene="yggN"
/function="orf; Unknown function"
/note="Residues 1 to 239 of 239 are 99.58 pct identical to residues 1 to 239 of 239 from Escherichia coli K-12 Strain MG1655: B2958"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58089.1"
/db_xref="GI:12517507"
/translation="MMRMLLAALSVTAMTAHAYQCSVTPRDDVIVSPQTVKGS NGNLVITPDGVMYNGQYSLNAAQROAKDYQAEALSTLPLWIDEGAKSVKARIAL DKIVQMGSSSKMRSLTKLDAOLKQSMNRIETRSNGLTEHYKALDQVXEQQLV NOAMGILQDSINEMGAKAVLKSGNPLQNVLSGLQSSIQIEWKKQKNDPOFGK DVCGRVTLDSRKALVGNLK"
/complement(2214..2570)
/gene="yggJ"
/note="synonym: Z4304"
/complement(2214..2570)
/gene="yggJ"
/function="orf; Unknown function"
/note="Residues 1 to 118 of 118 are 99.15 pct identical to residues 1 to 118 of 118 from Escherichia coli K-12 Strain MG1655: B2959"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58090.1"
/db_xref="GI:12517508"
/translation="MGLNVREGEIMAKNRRRLRKXKHIDEFOELGFSVAMRPFGTS BEQIDKTVDDFINEVIEPNKLAFDGSGYLAWGLICWQETGCTEBHQAIVRKLEER NLDEVRTSELFDYVWD"
/complement(2540..3259)
/gene="yggH"
/note="synonym: Z4305"
/complement(2540..3259)
/gene="yggH"
/function="orf; Unknown function"
/note="Residues 1 to 239 of 239 are 99.58 pct identical to residues 1 to 239 of 239 from Escherichia coli K-12 Strain MG1655: B2960"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58091.1"
/db_xref="GI:12517509"
/translation="WKNDVISPEFENGRLRIRSVRQGLTKGQHALENTWVP MGVTFSEMDLPALFGRAPVTLTIGFGMGSLVAMAKRFEQDFLGEVHSPGVGA CIASAHEGLSLRWCHDAVBLHKIIPDNSLRVWQVLPFPDPWHKRNKRRIYQVP FAEYKSLQLGCVFHMATDWEFPAEHMIEVMSSIDGYKNLSNDYVPRASRPVTK FEQRHRLGHGVNDLWFERVK"

```

```

gene
CDS

```

```

3420..4472
/gene="mutY"
/note="synonym: Z4306"
3420..4472
/gene="mutY"
/function="enzyme; Macromolecule synthesis, modification: DNA - replication, repair, restr./modific'n"
/note="Residues 1 to 350 of 350 are 99.42 pct identical to residues 1 to 350 of 350 from Escherichia coli K-12 Strain MG1655: B2961"
/codon_start=1
/transl_table=11
/product="adenine glycosylase; G.C --> T.A transversions"
/protein_id="AAG58092.1"
/db_xref="GI:12517510"
/translation="MQASQFSAQVLDVYDKYGRKTLPMQIDKTPYKVLSEVLMQOTQ VATVPIYFERMARFPVTDLANAPLDEVLHLWTGLGYARARNLHKAQQVATLHGG KFPETFEVALPGVGRSTAGAILSLSGKHFPIIDGNVNRVLRKCYAVSGWPGKKEV ENKLWSLSEQVTPAVGVERFQAQMDLGMICRSKPKCSLCPLQNGCIAATNNSWSL YPGKKPKQTLDPERTGYFLLQDEHDEVLIAQRPSPGLWGGLYCFQFQADESLRWLAQ RQIAADNLQTLTAFRHTFSHFHLDIVPMWLPVSSFTGCMDEGNALWYNLAQPPSVGLA APERLLQQLRTGAPV"
4500..4775
/gene="yggX"
/note="synonym: Z4307"
4500..4775
/gene="yggX"
/function="orf; Unknown function"
/note="Residues 1 to 91 of 91 are 100.00 pct identical to residues 1 to 91 of 91 from Escherichia coli K-12 Strain MG1655: B2962"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AAG58093.1"
/db_xref="GI:12517511"
/translation="MSRTIFCTFLQREAGQDPQLYPGELGKRIYNEISKEAWAQWH KOTMLNEKKLNNMNAEHRKLLEQEMVNFLEGEVHIEGTTPEDKK"
4837..5919
/gene="mltC"
/note="synonym: Z4308"
4837..5919
/gene="mltC"
/function="putative enzyme; Cell envelop: Murein sacculus, peptidoglycan"
/note="Residues 1 to 360 of 360 are 99.72 pct identical to residues 1 to 360 of 360 from Escherichia coli K-12 Strain MG1655: B2963"
/codon_start=1
/transl_table=11
/product="membrane-bound lytic murein transglycosylase C"
/protein_id="AAG58094.1"
/db_xref="GI:12517512"
/translation="MMKYLALALIAPLISCTTKKGTNEAWVKDNGFDILMQQ FAHNIEWGFKEVVIAGDPYKYDTQTRSHINFDGTTITETIAGTEPAHLRR AIITLLMGDDSPSDLYSDVDDITISKEPFLYGQVNDTGPRIWEGRASFADYLL KNRLKRSNGLRIYVVTINMVLNHLDKAHKYLGMVROASKYGVDESLIIMOTE SSENFPYAVSRSDALGDMOVQHTAGKDVFRSGKSGTSPRSKDFDPASNDTGTAYLA MLNNVILGIDNPTSRRTAVITAYNGAGSVLRVFSNDKIQAAANIINTMTPGDVYQTL TRHPSASRRYLYKVNTAQSKYRRR"
6073..7377
/gene="nupG"
/note="synonym: Z4309"
6073..7377
/gene="nupG"
/function="transport; Transport of small molecules: Nucleosides, purines, pyrimidines"
/note="Residues 1 to 434 of 434 are 99.76 pct identical to residues 1 to 434 of 434 from Escherichia coli K-12 Strain MG1655: B2964"
/codon_start=1
/transl_table=11
/product="transport of nucleosides, permease protein"

```

/protein\_id="AAG58095.1"  
/db\_xref="GI:12517513"  
/translation="MCFKHSASRCEENNNKLQKILSLFQCLWGLWLTLLGSY  
MFVTLKFDGASIGAVYSSGLIAAVFMPALLIGIVADKWLAKWVAICHTICAITLFWA  
AQVTPPAMELVILINSFAYMPTGLINTISYVELONAGMDIVDRPPIRWITGIFGI  
MAMVVSLSGFEUHQYIYGAAISAILVFLTLPHLPVAKQOAGNSWTLGLGLDAP  
ALFKKEMALFFIFSLMLGAEQLTNMFGNDFLSPKDFKMPASSFVQHASIIMSIS  
QISFTLILPPEFLSYGKINVMISIVAMIRFALFALPGDPTFPFGVLLVLSMIVY  
GCAPDFNIGSVFVEKEVSPAIRAQGMFLMNTGFGCILGIVSGKVVEMYTONG  
ITDQTVLILPAGYSVVLAFAMFMYKHVRVPTGTQTVSH"  
complement (7427..9622)

Query Match	26.2%; Score 296.8; DB 1; Length 10749;
Best Local Similarity	56.9%; Pred. No. 8.1e-74;
Matches 566; Conservative	0; Mismatches 422; Indels 6; Gaps 1;
QY	119 TGGCTTAAACCCCAAGTGACTATCTAGCCACAGGAGGACCATCGCTGTTTCGGGGGAAT 178
DB	1074 TGGCATTAACCAATATACCATTTTAGCAACCGGGGACCATTCGCCGTGCTGACT 1015
QY	179 CTAGCGTCAAGAGTAGTACTCTGTGTGAGCAGTACCGTGTGATAGTCTTTCGAGCG 238
DB	1014 CGCAACCAATCTAACTACACAGCGGTAAAGTTGGCGTAGAAAATCTGGTTAATGCGG 955
QY	239 TCCCTGCCATCAACGACCTAGCCCATCAAGGGTGAACAGATCTCAAGCATTTGCTCCC 298
DB	954 TGCCGCAACTGAAGACATTTGGAACGTTAAAGCGGAGCAGGTAGTGAATTCGGCTCCC 895
QY	299 AAGAGATACGGGTAAAGTGTGGCTTAAACTAGCCAGCGGTGTCAATGAGTCTTCGCC 358
DB	894 AGGACATGAACGATGATCTGTGCTGACACTGGCGGAAAAAATTAAC-----CACCGACT 841
QY	359 AAAAGAGACCGAAGCGGTGATCATCAACCATGGAACGTGACACCATGGAAGAGCGGTT 418
DB	840 GCGATAAAACCGACGCTTTCGTCATTAACCGTACCGGACGATGGAAGAAACCGCTT 781
QY	419 TCTTCTCAACCTCAGCGTGAAGGAAAGCCAAACCTCTGCTGCTGTAGCGCGCATCGCTC 478
DB	780 ACTTCTCGACCTGACGTTGAAATGGCAACCGGTGTGATGTTGCGCGCAATGCGTC 721
QY	479 CAGGCTCTTCCATGATGCTGTATGCGCCCATGAATCTTATACGCGGTGAATGTAGCGA 538
DB	720 CGTCCAGCTTATGAGCGCAGACGGTCCATTCAACCTGTATAACGCGGTAGTGACCGCAG 661
QY	539 TCACCAAGACCTCTACTACAAAGAGTGGTGTGATGATGAACGATGATGATTCAGCGCG 598
DB	660 CTGATAAAGCCTCCGCTAATCGTGGCGTGTGTTGATGAACGACACCGTACTGAGCG 601
QY	599 CCAGAGAGCGACCAAGCTCAACACACCGCGAGTCAATGCTTGTCTCGCCCAACACAG 658
DB	600 GTCCGATGTACCAAAACCAACACACCGAGCGTAGCGACCTTCAAGTCTGTTAACTAG 541
QY	659 GTAAATTCGGCAGCTATTATATGGCAAGTCAAGTATTTCACTCAATCCGTTCCGACTC 718
DB	540 GTCCGCTGGGATACATTCAACCGTAAAGTTGACTACCAACGTAACCCCGCGACGTAAGC 481
QY	719 ACACCTTTCGAGTGATTTGATTTAGCAAAATCGAAGAACCTCCCGACGATCGATATTC 778
DB	480 ACACAGGACACCGCGCTTCGATGCTCTTAAGCTGATGAACCTCCCGAAGTCCGATG 421
QY	779 TTTACGCTCAACCCCGATGATGATGTTTTAGTCAATGAGCGCCCTTCAGCGAGGACCA 838
DB	420 TTTATAACTAGCTAAGCATCCGATCTTCCGCGTAAAGCCCTGGTAGATGCGGGCTATG 361
QY	839 AAGCAATCATCATGACGAGCATGGGCAATGGGACCTTTCCCTTTTGACTCAAAATGCTC 898
DB	360 ATGGCATCGTGAAGCGCTGTGGGTAAACCGCAACCTGTATATAATCCGTTTTCACACGC 301
QY	899 TTGAAAGACGACCAAAATCAGGCGTGTGCTCGCTCGAAGCTCTAGAGTGGCGAGTGTT 958
DB	300 TGGGACCGCGGGAACAAACGGTACTGCTGCTGCTCTTCCCGGTACCGACGCGG 241
QY	959 CCACCAACCAAGAGGCTGAAGTGGATGATAAGAAACCTTGGTTTGTGGCTACAGAGTCT 1018

Db	240 CTACCACTAGGATCGCAAGTGGATGATCGAATAATCGGTTTATTGCTCTGGTACGC 181
QY	1019 TCACCCCTAAAAGCCAGAGTGTCTTTATGTTAGCTCCACCAAACTAGTATAGAG 1078
Db	180 TGAACCCGCAAAAAGCGCGGCTCTGCTCAACTGGCTCTGACGCAACCAAGATCCGC 121
QY	1079 AGGCGATCCCAAAAGATCTTCTCCACCTATTAAATC 1112
Db	120 AGCAGATCCACAGATCTTCAATCAGTACTAATC 87

RESULT 10	AP002563/c
LOCUS	266658 bp DNA linear BCT 07-MAR-2001
DEFINITION	Escherichia coli O157:H7 DNA, complete genome, section 14/20.
ACCESSION	AP002563 BA000007
VERSION	AP002563.1 GI:13363121
KEYWORDS	Escherichia coli O157:H7
SOURCE	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
ORGANISM	1 (sites)
REFERENCE	Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S., Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T., Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and Shinagawa, H.
AUTHORS	Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7 derived from the Sakai outbreak
TITLE	Genes Genet. Syst. 74 (5), 227-239 (1999)
JOURNAL	20198780
MEDLINE	10734605
PUBMED	2 (sites)
REFERENCE	Ohnishi, M., Murata, T., Nakayama, K., Kuhara, S., Hattori, M., Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shinagawa, H. and Hayashi, T.
AUTHORS	Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an Escherichia coli K-12 strain MG1655
TITLE	Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
JOURNAL	20557356
MEDLINE	11108008
PUBMED	3 (sites)
REFERENCE	Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S., Yutsudo, C.H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T., Yamamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and Shinagawa, H.
AUTHORS	Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli O157:H7 strain derived from the Sakai outbreak
TITLE	Gene 258 (1-2), 127-139 (2000)
JOURNAL	20564182
MEDLINE	11111050
PUBMED	4 (sites)
REFERENCE	Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogaawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. and Shinagawa, H.
AUTHORS	Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12
TITLE	DNA Res. 8 (1), 11-22 (2001)
JOURNAL	21156231
MEDLINE	11258796
PUBMED	5 (bases 1 to 266658)
REFERENCE	Ohnishi, M., Kurokawa, K., Makino, K., Yasunaga, T., Shinagawa, H. and Hayashi, T.
AUTHORS	Direct Submission
TITLE	Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail:ken@gen-info.osaka-u.ac.jp)
JOURNAL	

URL: <http://www.gen-info.osaka-u.ac.jp/>, Tel: 81-6-6879-8365,  
Fax: 81-6-6879-2047

COMMENT  
FEATURES

## source

## Location/Qualifiers

1. 266658  
/organism="Escherichia coli O157:H7"  
/mol\_type="genomic DNA"  
/strain="O157:H7"  
/sub\_strain="RIMD 0509952"  
/db\_xref="taxon:83334"  
complement(244..693)  
/gene="ECs3650"  
complement(244..693)  
/note="similar to B2790 ECOLI gi|1789154 percent identity  
100 in 149 aa (Conserved in E.coli K-12)"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB37073.1"  
/db\_xref="GI:13363122"  
/translation="MAEIGIFVTGWSLVAEAEAILTAQGHKATVPEDPEISDW  
LPYODKYVLVVTSTGGDLPSIVPLFGKIDSLGQPNLRYGVIALGDSYVNFPCN  
GSKQFDALLQSQARVGEMLLIDASENPEPESNPVVEQMGTLSS"  
complement(711..1493)  
/gene="ECs3651"  
complement(711..1493)  
/gene="ECs3651"  
/note="similar to YQCH ECOLI gi|1789155 percent identity  
99 in 260 aa (Conserved in E.coli K-12)"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB37074.1"  
/db\_xref="GI:13363123"  
/translation="MLEILYQDEWLVAVKNPSGMLVHRSWLDREBKVVVMQTVRQIG  
QHVTAHRLDRPTSGVLMLGSSSEAGRLLAQFQHOQIKRYHAIVRWLMBEAVLDY  
FLVEELDKADKFAREQPQPVAVHYRGLATVEMPATGYPTTGYGLVELEPKTR  
KHQRLHRLHRIIGDSKHGDLQRNKGAEHFGQLRLMLHASQLSLTHPTGPEPLT  
IHAGLDITWQALSQFQWRGLPENVERVEFSAPSQDGERSS"  
complement(1493..1822)  
/gene="ECs3652"  
complement(1493..1822)  
/note="similar to B2792 ECOLI gi|1789156 percent identity  
98 in 109 aa (Conserved in E.coli K-12)"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB37075.1"  
/db\_xref="GI:13363124"  
/translation="WTHDRVRLQLEALLREHQRWNRDEPOPHQFNSTQPFMDT  
MEPLWLOWLIPRMHLLNNQLPGFAVAPYVEALATDHPQRLIALLEKLDA  
LPADDAAS"  
complement(1865..2233)  
/gene="csrB"  
complement(1865..2233)  
/gene="csrB"  
/product="CsrB RNA"  
/note="inhibitor of CsrA"  
complement(2444..2989)  
/gene="ECs3653"  
complement(2444..2989)  
/note="similar to SYD ECOLI gi|1789157 percent identity 98  
in 181 aa (Conserved in E.coli K-12)"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="Syd protein"

## gene

## CDS

/protein\_id="BAB37076.1"  
/db\_xref="GI:13363125"  
/translation="MDDLTAQALDKFTARYCDAWHEEHKSWPLSELYGVSPSPCIIST  
TEDAVYWPQOPFTGEQNVNAVERAFDIVIOPTINTFTYTOAGDMHAQFGDIKLTLLQ  
TWSEDDPRRQENLIGHLVTKOKLKLPTLTFIATLEEBLEVISVCNLSGEVCKETLGT  
RKRTLHAPNLAEFLNQLKPLL"  
3057..3905  
/gene="ECs3654"  
3057..3905  
/note="similar to YQCD ECOLI gi|1789158 percent identity  
98 in 282 aa (Conserved in E.coli K-12)"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB37077.1"  
/db\_xref="GI:13363126"  
/translation="MSSYANHOALAGLTCKSTDYRDTYDASLLQGVPRSLNRDPLGL  
KADNLPQGTDIWTLVELSWLNAGLQVAVGHVELDYTSVNLLESKSKFLYLSNFQ  
TRNNWDEQVTLERDLSTCAQGEVSVALYRLDEGOPIGHFNGTCTDDQDITIDNY  
EFTDYLENATSEKVEEVLVSHLLSKNCLITHQPDWGSIOIYVRGQIOREKLLRY  
LVSRHNEHFECQVERIFENDLLRFCQPEKLSVYARYTRGGLDINPWRNSDFVPS  
TRLVRQ"  
4017..5381  
/gene="ECs3655"  
4017..5381  
/note="similar to YGDH ECOLI gi|1789159 percent identity  
99 in 454 aa (Conserved in E.coli K-12)"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BAB37078.1"  
/db\_xref="GI:13363127"  
/translation="MITHSPLGMDMLSQLEVDMLKRTASSDLQYLFNCSLAVLNS  
GSLTNSKELSRFENFDINLRRERGVKLELNPPEEAFVDGRIIRALQANLPAVL  
DILFVYQJHNTVRPNLNDLSVHLTNLVSILRNARALHVGSEPNVVCWGHSIN  
ENEYLARVGNQLGRLNLCITCGFGAEAPGAAGAAHQAYKDSRFIGNTEPS  
ITAAPVPLVNLLEIIMPDIKRLAEFVRIAGHIIIPPGVGVTABELLYLLGLMNP  
NKDQPLILITQKESADYFVLDFEVHTLGENARRHYRIIDDAEAVARQMKMWP  
LVKENRDTGDAYSFNWSMRIAPDLQMPPEPHENMANLKYPDQPVVEVLADLRRAF  
SGIVAGNVKEVGIRAEIEPQVYKINGDKIMRMDLLQGFVQAHRMKLPGSAVIPC  
EICT"  
5938..7227  
/gene="ECs3656"  
5938..7227  
/note="similar to SDAC ECOLI gi|1789160 percent identity  
100 in 429 aa (Conserved in E.coli K-12)"  
/codon\_start=1  
/evidence=not experimental  
/transl\_table=11  
/product="probable serine transporter"  
/protein\_id="BAB37079.1"  
/db\_xref="GI:13363128"  
/translation="METQTSTIASKDSRSARWKTDTMMMLGLYGTAGVGLFLPIN  
AGVGMIPLIIMAILAPMTFAHRLGTRFVLGKNGPDEITEVEEHFGIGAGKLIT  
LLYFPAIYPIILVYSVALTNVSESHQGMTPPPRILSLILIVGMNTIVRFGQM  
IVKAWSLVFPVGVMLALYLIPOWNGALETSLDTSATNGLMTLWLAIPVM  
VFSFNHSPITISSFAVAREEYGDMAEQCKSLIAFAHIMMVLTMFVFCVLSLTPA  
DLAAKEQNI SILSYLANHFNAPVIAWMAPIIAITATKPLGHYLGAREGFGNWK  
SLRGKGSIEINKLNRITVITVATNPISILGMIETLGGPIIAMILLFMPMY  
AIOKVPAMRYKSGHISNVVVVMGLIAISFYSLPS"  
7285..8652  
/gene="ECs3657"  
7285..8652

Query Match 26.2%; Score 296.8; DB 1; Length 266658;  
Best Local Similarity 56.9%; Pred. No. 1.4e-73;  
Matches 566; Conservative 0; Mismatches 422; Indels 6; Gaps 1;

QY	119	TGGCTTAACCCCAAGTGAAGTCTATCTAGCCACAGGAGGACCATCGCTGGTTCCGGGGAAT	178	ACCESSION	AE016766	AE014075
Db	197570	TGGCAATACCAATATACCATTTTACCAACCGCGGACCATTCGCGGTGGTGAAT	197511	VERSION	AE016766.1	GI:26109707
QY	179	CTAGCTCAAGTAGTACTCTGCTGGAGAGTACCGGTGATAAGCTTCTTTCACGCG	238	KEYWORDS	Escherichia coli CFT073	
Db	197510	CCGCAACCAATCTTAATACACAGCGGGTAAAGTTGGCGTAGAAAATCTGGTTAATCGG	197451	SOURCE	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	
QY	239	TCCCTGCCATCAAGCACTAGCAACCATCAAGGGTGAACAGATCTCAAGCATTCGGTCCC	298	ORGANISM	Enterobacteriaceae; Escherichia	
Db	197450	TCCCGCACTGAAGGACATTCGCAACGTTAAAGCGGAGCAGGTAGTGAATATCGCTCCC	197391	REFERENCE	1 (bases 1 to 303121)	
QY	299	AAGAGATGACGGGTAAAGTGGCTTAAACTAGCCAGCGGTCTCAATGAGTCTCTCGCC	358	AUTHORS	Welch, R.A., Burland, V., Plunkett, G.D. III, Redford, P., Roesch, P.,	
Db	197390	AGGACATCAACGATGATCTGGCTGACACTCGCGAATAAATTAAC-----CACCGACT	197337	JOURNAL	Rasko, D.A., Buckles, E.L., Liou, S.-R., Boutin, A., Hackett, J.,	
QY	359	AAAAGAGACGAGCGGTGATCATCCCATGGAATCGACACATGAAGAGACCGCTT	418	PUBMED	Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C.,	
Db	197336	CGGATAAACCGAGCGCTTCGTCTATCTACGTTACCGACACGATGAAGAACCGCTT	197277	REFERENCE	Perna, N.T., Mobley, H.L.T., Donnenberg, M.S. and Blattner, F.R.	
QY	419	TCTTCTCAACCTCACGCTGAAAGCCAAACCTGCTGCTGCTGTTAGGCGCATGCGTC	478	AUTHORS	Extensive Mosaic Structure Revealed by the Complete Genome Sequence	
Db	197276	ACTTCTCGACTGAGCTGGAATCGGACAAACCGGTGATGTCGCGCAATGCGTC	197217	TITLE	of Uropathogenic Escherichia coli	
QY	479	CAGGCTCTTCATGATGCTGATGCCCCATGATCTCTATAACCGCTGAATGAGCA	538	JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 17020-17024 (2002)	
Db	197216	CGTCCACGCTATGAGCGACGAGCGGTCCATTAACCTGTATACCGGTAGTACCGGAG	197157	PUBMED	12471157	
QY	539	TCAACAAAGCCTCTACTTAACAAAGAGTGGTGAATGATGAACGATGAGATTCACGCG	598	FEATURES	2 (bases 1 to 303121)	
Db	197156	CTGATAAAGCCTCCCGTATCTGTCGGTGTCTGTCGTGATGAACGACCGCTACTGGAG	197097	source	Welch, R.A., Burland, V., Plunkett, G.D. III, Redford, P., Roesch, P.,	
QY	599	CCAGAGACGCAACCAAGCTCAACACACCGCAAGTCAATGCAATTCGTCGCCCAACAG	658	gene	Rasko, D.A., Buckles, E.L., Liou, S.-R., Boutin, A., Hackett, J.,	
Db	197096	GTCCGATGTACCAAAACCAACACCGCAAGTCAATGCAATTCGTCGCCCAACAG	197037	CDS	Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C.,	
QY	659	GTAAATCGGCACAGTCTATTATGCAAGTCAAGTATTCATCAATCCGTTCCACCTTC	718	gene	Perna, N.T., Mobley, H.L.T., Donnenberg, M.S. and Blattner, F.R.	
Db	197036	GTCCCGTGGATACATTCACACGGTAAGATTGATCAACAGTACCCCGGACGTAAGC	196977	CDS	Direct Submission	
QY	719	ACACCTCTCAAGTGAAGTTGATATTAGCAAAATCGAAGATCCCGAGTCAATATTC	778	FEATURES	Submitted (20-JUN-2002) Genetics Laboratory, University of	
Db	196976	ACACGAGCAGACCGCGTTTCGATGTCTCTAAGCTGAATGAATCGCGAAGTCGGCATTG	196917	source	Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA	
QY	779	TTTACGCTCACCCGATGATATGATGTTTATGATGATGATGATGATGATGATGATGAT	838	Location/Qualifiers	1. 303121	
Db	196916	TTTATAACTACGCTAACGATCCGATCTTCCCGCTAAAGCCCTGGTAGATCGCGGTATG	196857	/organism="Escherichia coli CFT073"	1. 303121	
QY	839	AAGGAATATCCATGACGAGCATGGCAATGGGAACCCCTTTCCTTTCGACTCAAAATGCTC	898	/mol_type="genomic DNA"	/organism="Escherichia coli CFT073"	
Db	196856	ATGGCATCGTGACGCTGGTGGTTAAACCGCAACCTGTATAAATCCGTTTGAACAGC	196797	/strains="CFT073"	/mol_type="genomic DNA"	
QY	899	TTGAAAGACGACCAATACGCGTATGTCGCTCGAAGCTCTAGATGGGAGTGGTT	958	/db_xref="taxon:199310"	/strains="CFT073"	
Db	196796	TGCGGACCGCGCGGAAACCGTACTGCACTGCTGGTCTTCCCGGTACCGACGGCG	196737	/gene="ygfU"	/db_xref="taxon:199310"	
QY	959	CAACACCGGAGGCTGAATGATGAAGAACTGCTGTTGTTGCTTACAGAGATC	1018	/locus_tag="c3465"	/gene="ygfU"	
Db	196736	CTACCACTCAGGATGCGAAGTGGATGATGCGAATACGTTTATGCTCTGTTACGC	196677	/locus_tag="c3466"	/locus_tag="c3465"	
QY	1019	TCAACCTCAAAAGCCAGAGTGTCTTATGTTAGCCCTCACCAAACTAGTATAGAG	1078	/codon_start=1	/locus_tag="c3466"	
Db	196676	TGAACCGGCAAAAGCGCGCTCTGCTGCAACTGGCTCTGACCAACCAACCAAGATCCG	196617	/transl_table=1	/locus_tag="c3466"	
QY	1079	AGGCAATCAAAAGATCTTCTCCACCTATTAATC 1112		/product="Hypothetical protein"	/locus_tag="c3466"	
Db	196616	AGCAGATCCAGCAGATCTTCAATCAGTACTAATC 196583		/protein_id="AA081914.1"	/locus_tag="c3466"	
QY	11	RESULT 11		/db_xref="GI:26109709"	/locus_tag="c3466"	
LOCUS	AE016766/c	303121 bp DNA linear BCT 24-DEC-2002		/translation="MSNNNNLLMRAIFSLLLFTLSIPVISDCVAMAIERSFKYMM	/locus_tag="c3466"	
DEFINITION	Escherichia coli CFT073 section 12 of 18 of the complete genome.			LLF"	/locus_tag="c3466"	
				2004. .2552	/locus_tag="c3466"	
				2004. .2552	/locus_tag="c3466"	
				/EC_number="5.3.3.2"	/locus_tag="c3466"	
				/function="putative"	/locus_tag="c3466"	
				/note="Escherichia coli K-12 ortholog: b2889; Escherichia	/locus_tag="c3466"	
				coli O157:H7 ortholog: z4227"	/locus_tag="c3466"	



QY	359	AAAAAGAGACCGGAGCGGTGATCATCACCCTGGAACCTGACACCATGGAAGAGACCGGCTT	418			
Db	68968	GGGATAAAAGTACGGCTTCGTCATTTACCCACGGTACCGACAGTGAAGAAACCGGCTT	68909			
QY	419	TCCTTCCTCAACCTCAGCGTGAAGAGCAAAACCTGCTGCGCTTTAGCGCGCATCGGTC	478			
Db	68908	ACTTCTCGACCTGACCGTGAATGCGAACAACCGGTGCTGCTGCGCTGAATGCGCC	68849			
QY	479	CAGGCTTTCCATGAGTGTGATGCGCCCATGAATCTCTATAACGCGTGAATGATGACGA	538			
Db	68848	CGTCCACGCTTATGAGCGCAGACGGTCCATTCACCTGTATACGCGGTAGTACCGCAG	68789			
QY	539	TCAACAAAGCCTCTACTAAACAAAGAGTGTGATGTGATGAACGATGAGATTCACGCG	598			
Db	68788	CTGATAAAGCCTTCGCTTAATCGTGGGTACTGCTAGTGTGATGAACGACACCGCTTGTATG	68729			
QY	599	CCAGAGAGCGACCAAGCTCAACACACCGCAGTCAATGCTTGTGCGCCCAACACAG	658			
Db	68728	GCGGTGATGTACCAAAACCAACACACCGCATGTAGCGACCTTCAAGTCTGTTAACTACG	68669			
QY	659	GTAAATCGGCACAGTCTATTATGCGAAAGTCGAGTATTTCATCAATCGCTTCGACCTC	718			
Db	68668	GTCTCTGGGTAACTTCAACACGTTAAGATTGACTACCAACGTTACCCGCGCAGTAAAGC	68609			
QY	719	ACACCTTTGCAAGTGTGATGATTAAGCAAAATCGAAGAACTCCCGACAGTGCATATTC	778			
Db	68608	ACACGAGGACACGCGGTTGATGCTCTAAGCTGATGAATGAACTCCCGAAGTCGGCAATTG	68549			
QY	779	TTTACGCTACCCCGATGATGATGTTTATGTCATGATGAGCGCTTCAGCGCAGGAGCCA	838			
Db	68548	TTTATAACTAGCTTAACGATCCGATCTTCGCGTAAAGCACTGTTATATAACCGTATTGACACCC	68429			
QY	839	RAGGAATCATCATCGACGATGGCAATGGGAACCCCTTTCCCTTTGACTCAAAATGCTC	898			
Db	68488	ATGGCATGTTAGCGCTGCGGTGGTAAACGCGCAACCTGTATATAACCGTATTGACACCC	68429			
QY	899	TTGAAAGAGCAGCCAAATCAGCGGTAGTCGTCGCTGGAAGCTCTAGAGTGGCAGTGGTT	958			
Db	68428	TTGCAACCGCTCGGAAAGCGCACTGCAAGTGTGCTTTCTCCCGGTACCGACGGCG	68369			
QY	959	CCACACCAACGAGGCTGAAGTGGATGATTAAGAACTTGTGTTTGTGGCTACAGAGATC	1018			
Db	68368	CTACCACTCAGATGCGGAATGGATGATGCGAATACGGCTTCGTCGCTCCCGCATGT	68309			
QY	1019	TCAACCTCAAAAGCAGAGTGTCTTATGTTAGTACCTCACCACCACTAGTATAGAG	1078			
Db	68308	TGAACCGCAAAAGCGCGGTTCTGCTGCAACTGCTCTGACGCAACTAAAGATCCG	68249			
QY	1079	AGCGATCCAAAGATCTTCTCCACCTATTAATC	1112			
Db	68248	AGCAGATCCAGCAGATCTTCAATCAGTACTAATC	68215			
RESULT 12						
U32758						
LOCUS	10173 bp DNA linear BCT 29-MAY-1998					
DEFINITION	Haemophilus influenzae Rd section 73 of 163 of the complete genome.					
ACCESSION	U32758 L42023					
VERSION	U32758.1 GI:1573747					
KEYWORDS	Haemophilus influenzae Rd					
SOURCE	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.					
ORGANISM	1 (bases 1 to 10173)					
REFERENCE	Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G.G., FitzHugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M.,					
AUTHORS	Smith, H.O. and Venter, J.C.					
Whole-genome random sequencing and assembly of Haemophilus influenzae Rd						
Science 269 (5223), 496-512 (1995)						
95350630						
7542800						
2 (bases 1 to 10173)						
Tatuzov, R.L., Mushagian, A.R., Bork, P., Brown, N.P., Hayes, W.S., Borodovsky, M., Rudd, K.E. and Koonin, E.V.						
Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli						
Curr. Biol. 6 (3), 279-291 (1996)						
96398784						
MEDLINE						
PUBMED						
8805245						
3 (bases 1 to 10173)						
White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.						
Direct Submission						
Submitted (25-JUL-1995) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA						
4 (bases 1 to 10173)						
White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.						
Direct Submission						
Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA						
The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatuzov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes						
5 (bases 1 to 10173)						
White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.						
Direct Submission						
Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA						
The whole genome was shifted by 588 nucleotides for a new start						
On Sep 30, 1996 this sequence version replaced gi:1221457.						
Location/Qualifiers						
1..10173						
/organism="Haemophilus influenzae Rd"						
/mol_type="genomic DNA"						
/db_xref="taxon:71421"						
247..1704						
/gene="HI0740"						
247..1704						
/genes="HI0740"						
/note="similar to GB:M34393 SP:PI8159 PID:142994						
PID:226139 GB:AL009126 percent identity: 38.72;						
identified by sequence similarity; putative"						
/codon_start=1						
/transl_table=11						
/product="phosphomannomutase (yhxB)"						
/protein_id="AAC22400.1"						
/db_xref="GI:1573748"						
/translation="MGNNRVLSQAAGGLAEYLKGYDKPSIVIGVDKRNKSDVFARD						
TAETMAGAGKAYLLPRKLPVLAIOYFTDAGVMTATNPPNDNGYKVLGKA						
NGGQIVSPADKIDIAALDKVAAGIOLPRSDNYVINDVDYATTKTSLAKEPA						
NTPGIVTAMHGVGYEVLKTLAKAGLPQPHVADQVMPDGTFTNPNPEEKGLD						
LAIVAKENAEFIANDPADRLAVAPDAQGNKSLHGNVVCFLGVLAKYQYQK						
QGTACLSVSPALAEIAKYSFQSEETLIGFKYIGKVGILGFEALGYLVDPKV						
RDKDGI3AAIVFLDLVRLKKQKTLADYAEFTKEFGVAVSGQISIRVSDLSIGKL						
MTALRNPPAEIAGVQAQFIDHIKTDQSDILVFNLENGRLIARFSGTEPKIFYL						
DARGKDPKADRVLAIEFDEGVRHILRQDAYKQDC"						
1853..2182						
/gene="HI0742"						
1853..2182						
/genes="HI0742"						
/note="hypothetical protein; identified by GeneMark;						
putative"						
/codon_start=1						
/transl_table=11						
/product="H. influenzae predicted coding region HI0742"						



Db	102782	AAATCTCGTATTAAGCTGGACAATTAAGTATTGATCTTTAAATTTGAAGCTGTACCAGA	102841
Qy	246	CAATCAACGACCTAGCCACCATCAAGGCTGAAAGAGATCTCAAGCATTTGGCTCCCAAGAGAT	305
Db	102842	AATGAAAAATATTGCGCAACATTAAGGTTGAGCAAAATTTGAAAAATAGTTTCAAGACAT	102901
Qy	306	GACGGTAAAGTGTGGCTTAAACTAGCCAAAGCGTGTCAATAGCTCTCCGCCAAAAAGA	365
Db	102902	GAATGACGAAGTCTGGCTAAACTTGGCAAAAGCCATCAATGCTC-----AATGTAAG	102955
Qy	366	GACCGAAGCGGTGATCATCAACCCATGGAATGACACATGAGAGAGACCGCTTCTTCTCT	425
Db	102956	TACTGATGATTTGTCATTTACCCATGATACAGATACCATGGAAGAGACGGCTTATTCTT	103015
Qy	426	CAACCTCACGCTGAAAGCCAAAAACCTGTCCTGCTTTAGAGCGCCATGCGTCCAGGCTC	485
Db	103016	AGATTTAACCGTAAATGTGAAAAACCGGTTGTTCTCGTTGGGCAATGCTCTGCAAC	103075
Qy	486	TTCCATGAGTGTGATGGCCCATGAAATCTCTATAAGCGCGTGAATGTACGATCAACAA	545
Db	103076	AGAAAAAGTCTGATGGCCCATTTAAATCTTTTAAATGCTGTCGTTGTCGAGCAGACAA	103135
Qy	546	AGCCTCTACTAACAAAGAGTGGTGAATGTGATGAACGATGAGATTCACGCCGCCAGAG	605
Db	103136	AAAAATCAAGTGTGCTGTTTGTAGTCGCAATGAAATGAAGTACTAGTGTCTCGCA	103195
Qy	606	AGCGACCAAGCTCAACACCCAGTCAATGCAATTTGCTTCGCCCAACACAGCTAAAT	665
Db	103196	TGTAACAAAAACAGTACGACCGCAGTCAAACTTCATTTCAACAAATTAATGTTCTCT	103255
Qy	666	CGGCAAGCTTATTTATGCAAGTGGTATTTCACTCAATCCGTTCCAGCTCAACCT	725
Db	103256	AGGCTATATTATACACGAAGTGGATGAAAGTTCGCCAGAAAGCAACATATACAT	103315
Qy	726	TGCAAGTGAATTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTGCATTTCTTTACG	785
Db	103316	CAACACTCCGTTTAACTAGTAAATATTAGTAGCTTACCCAAAGTGGGATTTATTATGC	103375
Qy	786	TCACCCGATGATATGATGTTTATGTTTATGCAATGCAAGCTTCCAGGAGGAGGCAAG	845
Db	103376	TTATTCMAATGCACTGTGCAACCACTTAAACGATTAATGCTGACACTTAGATCGTTTAA	103435
Qy	846	CATCCATGAGGCTGGGCAATGGGAACTTCCCTTTGACTCAAAATGCTCTTGAATA	905
Db	103436	TGATCTGAGGAGTTGGCAATGGAAATGTTAATGCTGACACTTAGATCGTTTAAATA	103495
Qy	906	AGCAGCAAAATCAGGCTGATGCTGCTGCAAGCTTACAGTGGGAGGCTTCCACAC	965
Db	103496	AGCGCAAAAGATAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	103555
Qy	966	CAAGAGGCTGAGTGTGATGATAAGAACTTGGTTTGTGGCTTACAGAGTCTCAACCC	1025
Db	103556	TCGTGAGCTGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG	103615
Qy	1026	TCAAAAAGCCAGAGTCTCTTATGTTAGTGGCTTCCCAAACTAGTGTAGAGAGCGAT	1085
Db	103616	ACAGAAAGCCGCGTCTCTTGCATTTAGCTTTTAACTCAAACTAAAGATCCNAAAT	103675
Qy	1086	CAAAAGATCTTCTCCACCTATTAA	1110
Db	103676	TCACAAATATTTGGAAGACTTCTAA	103700

RESULT 14	AR274513_08	Sequence split into 19 fragments		LOCUS AR274513	Accession AR274513
WPCOMMENT		Fragment Name	Begin	End	
		AR274513_00	1	110000	
		AR274513_01	100001	210000	
		AR274513_02	200001	310000	
		AR274513_03	300001	410000	
		AR274513_04	400001	510000	

Qy	606	AGCGACCAAGCTCAACACCGCAGTCAATGCAATTTGCTTGGCCCAACACAGCTAAAT	665
Db	3989	TGTAACAAAAACAGTACGACCGCAGTGCAAAGTTCATTTCAACCAATTTATGTTCTCT	4048
Qy	666	CGGCAAGCTTATTTATGCAAAAGTGCAGTATTTCACTCAATCCGTTCCAGCTCAACCT	725
Db	4049	AGGCTATATTATACAGCAAGTGGACTATGAAGCTTCCCGCAAGCAACATACAT	4108
Qy	726	TGCAAGTGAATTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTGCATATTCTTTAGC	785
Db	4109	CAACACTCCGTTTAACTAGTAAATTTAGTAGCCTTACCCAAAGTGGGATTTATTATGC	4168
Qy	786	TCACCCGATGATGATGATGATTTAGTCAATGAGCCCTTCCAGGAGGCGCAAGGAAT	845
Db	4169	TTATTCAAATGACCTGTGCAACCAATTAACGCAATTTACTCAATGCTGGCTATCAAGGAT	4228
Qy	846	CAATCCATGAGGATGGGCAATGGGAACCTTTCCTTTGACTCAAAATGCTCTTTGAAAA	905
Db	4229	TGATCTGAGGAGTTGGCAATGGAAATGTTAATGCTGACACTTAGATCGCTTAGAAAA	4288
Qy	906	AGCAGCAATCAGGCTAGTCTGCTGCAAGCTCTAGAGTGGGCGAGTGGTTCACACAC	965
Db	4289	AGCGCAAAAGATAGCTGCTGTTAGTCCGTTCTTCTGCTGATACCAACCGGTTATACAC	4348
Qy	966	CCAAGAGGCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1025
Db	4349	TCGTTGACGCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG	4408
Qy	1026	TCAAAAAGCAGAGTCTTCTTATGTTAGCCCTCACCAAACTAGTGTAGAGAGGCGAT	1085
Db	4409	ACAGAAAGCCGCGTCTCTGCAATTTAGCTTTAACTCAAACTAAAGATCCNAAAT	4468
Qy	1086	CAAAAGATCTTCTCCACCTATTAA	1110
Db	4469	TCACAAATTTTGAAGACTTCTAA	4493

RESULT 13  
AR274513\_07  
WPCOMMENT  
Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000
AR274513_05	500001	610000
AR274513_06	600001	710000
AR274513_07	700001	810000
AR274513_08	800001	910000
AR274513_09	900001	1010000
AR274513_10	1000001	1110000
AR274513_11	1100001	1210000
AR274513_12	1200001	1310000
AR274513_13	1300001	1410000
AR274513_14	1400001	1510000
AR274513_15	1500001	1610000
AR274513_16	1600001	1710000
AR274513_17	1700001	1810000
AR274513_18	1800001	1830121

Continuation (8 of 19) of AR274513 from base 700001 (AR274513 Sequence 1 from patent US

Query Match	26.2%	Score 296.4	DB 6	Length 110000
Best Local Similarity	57.1%	Pred. No. 1.5e-73		
Matches	562	Conservative	0	Mismatches 417; Indels 6; Gaps 1;
Qy	126	ACCCCAAGTGAATCTTACGACAGAGGCAACATGCTGGTTCGGGGAATCTAGCT	185	
Db	102722	ACCAATATTACAACTTTGGCAACGGGTGTTACATTTGCAAGAGCGGCAAGTTCCGT	102781	
Qy	186	CAAGAGTAGTACTCTCTGAGGAGTCAACCGTTGATAAGCTTCTTTCAGCCGCTCCCTGC	245	

Query Match	26.1%	Score 295.2	DB 1	Length 1530
Best Local Similarity	56.8%	Pred. No. 1.7e-73		
Matches 565	Conservative 0	Mismatches 423	Indels 6	Gaps 1

119	TGGCTAAACCCCAAGTCGACTATCTACGCACAGAGGCACATCGCTGGTTCGGGGGAAT	178
446	TGGCATTACCAATATACCACTTTTGCACACGGCGGGACCAATGCGCGTGGTGGTGACT	505

Thu Aug 28 09:00:59 2003

179	CTAGCGTCAAGAGTAGCTACTCTCTCTGGAGCAGTCAACGGTTGATGAAGCTTCTTTGAGCGCG	238
506	CGCMAACCAATCTAACTACACAGTGGGTAAAGTTGGCGGTAGAAAATCTGGTTAATGCGG	565
239	TCCCTGCCATCAACGACCTTAGCCACATCAAGGTTGAACAGATCTCAAGCATTTGGCTCC	298
566	TGCCCAACTAAAGACATTGCGAACGTTTAAGCGGAGCAGGTAGTGAATATCGGCTCC	625
299	AAGACATGACCGGTAAAGTGTGGCTTAAACTAGCCAAAGCGTGTCAATAGACTCCTCGCCG	358
626	AGGACATGAACGATTAATGTCTGGCTGACACTGGCGGAAAAAATTAA-----CACCGACT	679
359	AAAAAGAGACCGAAGCCGTGATCATACCCATGGAATGACACCATGGAAGAGACCGCTT	418
680	CGGTAAGACCGACGGCTTGGTCAATTACCCACGGTACCGACACGATGGAAGAACTGCTT	739
419	TCTTCTCAACCTCAGGTTGAAGCCAAAACCTGTCTGCTTTGTAGGCGCCATCGCTC	478
740	ACTTCTCGACCTGACGGTGAATCGCAAAACCGGTGGTGAATGGTGGCGCATGGGTC	799
479	CAGGCTCTTCCATGAGTGTGATGGCCCATGAATCTCTATAACCGCGTGAATGTACGGA	538
800	CGTCCACGCTATGACGCGACAGCGTCCATTCAACCTGTATTAACGCGGTAGTACCAG	859
539	TCAACAAAGCCTCTACTATAAAGAGAGTGGTGATTTGTATGAACGATGAGATTCACGCCG	598
860	CTGATAAAGCCTCCGCCAACCGTGGCGTCTGTGTAGTGAATGACACCGTGTGTGATG	919
599	CGAGAGAGCGACCAAGCTTCMACCACCACGCGAGTCAATGCAATTTGCTTCGCCCAACACAG	658
920	GCGGTGAGTGCACCAAAACCAACACCGAGCGTCAATGAGTCTGTTAACTACG	979
659	GTAATAATCGCACAGTCTATTATGCGAAAAGTCGAGTATTTACTCAATCCGGTTGCACTC	718
980	GTCTCTGGGTACTATTCACACGTAAGATTGACTACCAGCGTACCCCGCAGCAGTAAGC	1039
719	ACACCTTTGCAAGTGAAGTTTGTATTAATAGCAAAATCGAAGAACTCCCGAGGTGATATTC	778
1040	ATACCAGCGACACGCCATTGCGTGTCTTAAGCTGAATGAATCTGCGCAAAAGTCGCAATG	1099
779	TTTACGCTACCCCGATGATGATGTTTGTAGTCAATGACGCGCTTCAGGCAGGAGCCA	838
1100	TTTATACTACGCTTAACGATCCGATCTTCGGCTTAAGACACTGGTATGTCGGGTATG	1159
839	AAGGAATCATCCATCGAGCATGGCAATCGGAACCCCTTTCCCTTTGACTCAAAATGCTC	898
1160	ATGGCATCTGTAGCGTGTGTGGTTAACGCAACCTGTATAAATCTGTGTTGCAACGC	1219
899	TTGAAAAAGACGCCAATCAGCGTAGTCTGTCGTCGAAGCTCTAGAGTGGGCAAGTGT	958
1220	TGGCCACCGCGCGGAAAACCGGTACTGCGAGTGTGGCGTCTTCCGCGTACCGACGGCG	1279
959	CCACCACCCACAGAGCTCAAGTGAATGATAGAAACTTGGTTTTGTGGCTACAGAGTCT	1018
1280	CTACCACTCAGGATCGCGAGTGGATGATGCGAAATACGCGTTCGTCGCTCTGGCACGC	1339
1019	TCAACCCCTCAAAAAGCCAGAGTCTTATGTTAGCCCTCACCAAACTAGTGTATAGAG	1078
1340	TGAAACCGCAAAAAGCGCGGCTCTGCTGCAACTGGCTCTGACGCAAAACAAAGATCCCG	1399
1079	AGCGGATCAAAAGATCTTCTCCACCTAATATC	1112
1400	AGCAGATCCAGAGATCTTCAATCAGTACTAATC	1433

Search completed: August 27, 2003, 02:34:29  
Job time : 4830.26 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:28:23 ; Search time 441.315 Seconds  
(without alignments)  
6930.345 Million cell updates/sec

Title: US-09-937-982-3  
Perfect score: 1133  
Sequence: 1 atggcagcagcatcatca.....aagaagggaattcttcc 1133

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 19Jun03.\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1129.8	99.7	1133	20	AA03474 Wolinella succinog
2	1129.8	99.7	1133	21	AAC62512 W. succinogenes as
3	300.4	26.5	1044	21	AA082834 L-asparaginase DNA
4	296.4	26.2	1830	17	AAT42063 Haemophilus influ
5	294	25.9	1848	18	AAT96346 Chimeric gene cont
6	259.2	22.9	1196	19	AA143328 H. pylori GHPO 187
7	223.2	19.7	2837	8	AA070557 Sequence encoding
8	215.8	19.0	1014	15	AAQ68438 Pseudomonas glutam

9	199	17.6	1174	25	ABZ80801 Erwinia carotovora
10	108.8	9.6	345	24	ABX66196 Helicobacter pylor
11	103.6	9.1	543	24	AA521124 cDNA encoding Tm 3
12	102	9.0	543	24	AA521116 cDNA encoding Tm 2
13	102	9.0	543	24	AA521118 cDNA encoding Tm 2
14	102	9.0	543	24	AA521122 cDNA encoding Tm 3
15	102	9.0	543	24	AA521126 cDNA encoding Tm 7
16	101.6	9.0	1436	25	ACC44571 pNOV4836 glucose i
17	101.4	8.9	552	19	AAV13947 Salmonella sefa ge
18	100.4	8.9	681	24	AA521115 cDNA encoding Tm 2
19	100.4	8.9	681	24	AA521121 cDNA encoding Tm 3
20	100.4	8.9	681	24	AA521125 cDNA encoding Tm 7
21	100.4	8.9	682	24	AA521117 cDNA encoding Tm 2
22	100.4	8.9	682	24	AA521123 cDNA encoding Tm 3
23	100.4	8.9	776	24	AA521119 cDNA encoding Tm 1
24	98.4	8.7	1435	25	ACC44570 pNOV4835 glucose i
25	96.4	8.5	543	24	AA521120 cDNA encoding Tm 1
26	94.2	8.3	900	24	ABV73014 Recombinant SOWp5
27	93.4	8.2	1247	20	AA529725 S. pneumoniae type
28	82.8	7.3	960	25	ABX07792 S. pneumoniae type
29	82.8	7.3	2162598	25	AB556454 Streptococcus pneu
30	79.8	7.0	316	25	ABZ80804 Erwinia carotovora
31	75.8	6.7	2365589	24	ABA90521 Genomic sequence o
32	72.6	6.5	2231	19	AAV52398 Streptococcus pneu
33	71.4	6.4	366	22	AAF25008 Nucleotide sequenc
34	71.4	6.3	969	22	AAH53259 S. epidermidis ope
35	71.4	6.3	981	24	ABN90934 Staphylococcus epi
36	71.4	6.3	4020	22	AAH54070 S. epidermidis gen
37	71	6.3	360	25	ABZ80805 Erwinia chrysanth
38	70.4	6.2	1230	22	AAF25127 Nucleotide sequenc
39	68.4	6.0	6682	20	AAH12993 Enterococcus faeca
40	68.4	6.0	6682	24	AB598788 Enterococcus faeca
41	67.6	6.0	619	21	AA250042 DNA encoding Hepat
42	67.4	5.9	654	21	AA250043 DNA encoding hepat
43	66.6	5.9	501	22	AA06207 Staphylococcus aur
44	66.2	5.8	1455	22	AA500250 LFn-Bcl-Xl apoptos
45	66.2	5.8	2847	22	AAF25014 Nucleotide sequenc

## ALIGNMENTS

RESULT 1  
AA03474  
ID AA03474 standard; DNA; 1133 BP.

AC AA03474;

XX 30-APR-1999 (first entry)

DE Wolinella succinogenes L-asparaginase DNA.

XX L-asparaginase; amplification; treatment; disease; asparagine depletion;  
KW malignant disease; haematology; lymphoma; leukaemia; myeloma; AIDS;  
KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;  
KW covalent modification; acylation; pharmacokinetic; immunogenic; spleen;  
KW hypersensitive; reduced toxicity; immunosuppression; allergy; thymus;  
KW lymphocyte marker distribution; hepatotoxicity; ss.

XX Wolinella succinogenes.

OS WO9856410-A1.

PN 17-DEC-1998.

XX 09-JUN-1998; 98WO-US11905.

PR 09-JUN-1997; 97US-0049085.

XX (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.

XX Durden DL;

XX



[illegible]

Db 62 TGGCATTACCAATATCACATTTTAGCAACCGCGGACCATTTGCCGGTGGTGACT 121  
 Qy 179 CTAGCGTCAAGAGTAGTCTCTGCTGAGAGTACACCGTTGATAAGCTTTTGACACCG 238  
 Db 122 CCGCAACCAATCTAATACACAGCGGTAAAGTTGGCTAGAGAAATCTGTTAATCGG 181  
 Qy 239 TCCTGTCATCAAGAGTCTAGCAGCAATCAAGGTGAACAGATCTCAAGCATTTGGTCCC 298  
 Db 182 TGGCGCAACTAAAGACATTTGCGAAGCTTAAAGCGAGCAGGTAGTGAATATCGGTCCC 241  
 Qy 299 AAGAGATGACGGTAAAGTGTGGCTTAACTAGCAAGCGTGTCAATGAGTCTCTCGCCC 358  
 Db 242 AGACATGAACAGTAATGTCTGCTGACACTGGCGAATAAATAA-----CACGACT 295  
 Qy 359 AAAAAAGAGACCGAAGCGGTGATCATCAACCCATGGAATGACACCATGGAAGAGACCGTT 418  
 Db 296 GCGATAAAGACCGAGCTTCTGCTATTAACCCAGCGTACCGACAGTGAAGAAACCGTT 355  
 Qy 419 TCTTCTCAACCTCAGGTGAAAGCAAAACCTGTCTGCTGTTAGCGGCCATCGTC 478  
 Db 356 ACTTCTGACCTGACGCTGAAATGCGAATGCAAAACCGGTGATGTTGCGGCAATCGGC 415  
 Qy 479 CAGGCTCTTCCATGAGTGTGTCGCCCATGATCTTATACCGCGTGAATGTAGCA 538  
 Db 416 CTTCCAGTCTTATGAGCGAGAGCGTCCATTAACCTGTATACCGCGTGTAGTACCGCAG 475  
 Qy 539 TCACAAAGCTCTACTAAACAAAGAGTGTGATTTGATGAACGATGATGATTCACGCG 598  
 Db 476 CTGATAAAGCTTCCGCTAATCTGCGGTGCTGTTGATGAACACACCGTACTGAGC 535  
 Qy 599 CCAGAGAGCGACCAAGCTCAACACACCGCAGTCAATGCTGCTGCTGTTAGCGGCCATCGTC 478  
 Db 536 GTCGCGATGTACCAAAACCAACACACCGCAGTACCGACCTTCAAGTCTGTTAACTAG 595  
 Qy 659 GFAAATCGGCACAGTCTATTATGCAAGTCTGATTTTCACTCAATCCGTTGACCTC 718  
 Db 596 GTCCTCTGGATACATTACAAACGTAAGTGTACTCAACGTTACCCCGGACGTAAGC 655  
 Qy 719 ACACCTTGCAAGTGTGATTTAGTAAATGCAAAAGTCTCCCGAGTTCGATATTC 778  
 Db 656 ACACAGGATACGCAATTCGATGTTCTAAGCTGAATGAGTGTGCGAAGTCTGCGATC 715  
 Qy 779 TTTACGCTCACCGGATGATCTGATGTTTGTAGTCAATGACGCTTACGAGCAGGCA 838  
 Db 716 TTTATACTAGCTAACGCTCCATCTTCCGGCTAAAGCACTGGTAGTCCGGCTATG 775  
 Qy 839 AAGCAATCATTCAGGCAATGGCAATGGAAACCTTTTCCCTTTGACTCAAAATGCTC 898  
 Db 776 ATGGCATCGTTAGCGCTGGTGGTAAATGTTAACTGTTAATCGTGTTCGACACCC 835  
 Qy 899 TTGAAAAGCAGCAAAATCAGGCTAGTCTGCTCGCTCGAAGCTCTAGAGTGGCAGTGGTT 958  
 Db 836 TGGCAACCGCGGAAAACGGCACTGAGTAGTGGCTTTCTTCCCGGTACCGACGGGTG 895  
 Qy 959 CACACCCCAAGAGGTGAAGTGAATGAAGAACTTGTGTTTGTGCTTACAGAGATC 1018  
 Db 896 CTACCACTCAGGATGCTGAAGTGAATGATGCGAAATACGCTTCTGCTGCTCTGGCAGC 955  
 Qy 1019 TCAACCTCAAAAGCAGAGTCTTCTTATGTTAGCCCTTACAAATCTAGTGTAGAG 1078  
 Db 956 TGAACCGGCAAAAGCGCGGTCTGCTGAGTGGCTCTGACGCAAAACCAAGATCCGC 1015  
 Qy 1079 AGCGATCCAAAGATCTTC 1098  
 Db 1016 AGCAGATCCAGCATCTTC 1035

RESULT 4  
 AAT42063  
 ID AAT42063 standard; DNA; 1830121 BP.  
 XX  
 AC AAT42063;

14-SEP-1999 (first entry)  
 Haemophilus influenzae complete genome sequence.  
 Genome; bacterium; Haemophilus influenzae; computer readable medium;  
 expression modulating fragment; regulation; gene expression; vector;  
 organism; open reading frame; ORF; ds.  
 Haemophilus influenzae.  
 WO9633276-A1.  
 24-OCT-1996.  
 22-APR-1996; 96WO-US05320.  
 07-JUN-1995; 95US-0487429.  
 21-APR-1995; 95US-0426787.  
 07-JUN-1995; 95US-0476102.  
 (HUMA-) HUMAN GENOME SCI INC.  
 (UYJO) UNIV JOHNS HOPKINS.  
 Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;  
 WPI; 1996-485782/48.  
 Haemophilus influenzae Rd genome recorded on computer readable  
 medium - useful for identifying commercially important nucleic acid  
 fragments by homology searching  
 Claim 1; Page 77.2-77.1091; 1291pp; English.  
 This sequence represents the complete genome sequence of the bacterium  
 Haemophilus influenzae strain Rd. The invention relates to a computer  
 readable medium (CRM) having recorded upon it the complete H. influenzae  
 nucleotide sequence (1), a representative fragment of (1) or a nucleotide  
 sequence at least 99% identical to (1). By providing the full-length  
 genomic sequence in a computer readable form, it is possible to identify  
 commercially important nucleic acid fragments and expression modulating  
 fragments (EMFs) of the Haemophilus genome. The EMFs can be used to  
 regulate the expression of a nucleic acid molecule. Vectors and altered  
 organisms comprising the predicted ORFs can be used to produce any of the  
 polypeptide fragments of the H. influenzae Rd genome.  
 Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;  
 Query Match 26.2%; Score 296.4; DB 17; Length 1830121;  
 Best Local Similarity 57.1%; Pred. No. 8.1e-81;  
 Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;  
 Qy 126 ACCCAAGTGACTATCTCTAGCCACAGAGCAGCATCGTGTTCGGGGAATCTAGCGT 185  
 Db 802722 ACCAAATATTACAAATCTTGCAACGGGTGGTACCATTTGAGGAAGCGGCAAGTTCGGT 802781  
 Qy 186 CAAGAGTAGTACTCTGCTGGAGCAGTCCCGTTGATAAGCTTCTTGCAGCCCTCCCTGC 245  
 Db 802782 AAATCTGCGTATAAGCTTGGACATTAAGTATTGATCTTAATTAAGAGCTGTACAGA 802841  
 Qy 245 CATCAACGACCTAGCCACCATCAAGGTTGAAAGAGATCTCAAGCATTTGGTCCCAAGAGAT 305  
 Db 802842 AATGAAAAATATTGCGCAACATTTAAAGGTGAGCAAAATGTAATAATAGTTTCAACAGCAT 802901  
 Qy 306 GACGGGTAAAGTGTGCTTAAACTAGCCAAAGCGTGTCAATGAGTCTCTGCCCAAAAGA 365  
 Db 802902 GAATGACGAGTCTGCTTAAACTGCAAAAGCCATCAATGCTC-----AATGTAAAG 802955  
 Qy 366 GACCGAAGCGGTGATCATCACCCATGGAACTGACACCATGGAAGAGACCGCTTCTTCTCT 425  
 Db 802956 TACTGATGGAATTTGTCAATTAACCATGTTACAGATACCATGGAAGAGACGGCTTATTCTT 803015  
 Qy 426 CAACCTCACGGTGAAAGACCAAAACCTGTCTGCTTGTAGCGGCATGCTCCAGGCTC 485

21-JUN-1988; 88US-0205748.  
31-AUG-1992; 92US-0938505.  
23-MAY-1995; 95US-0447422.  
(HYBR-) HYBRISENS LTD.  
Ramjeesingh M, Rothstein A, Shami EY;  
WPI; 1997-558200/51.  
Self-protecting chimeric polypeptide comprising biologically active sequence and single-chain antibody sequence - has resistance to e.g. disrupting temperature, presence of proteolytic enzymes, etc.  
Example 2; Columns 27-30; 29pp; English.  
The present sequence is a chimeric gene containing the cDNA for the immunoglobulin (Ig) heavy and light chain variable regions of a murine anti-asparaginase II monoclonal antibody (MAB), and human asparaginase II. The gene was used in the preparation of a novel recombinant chimeric polypeptide, comprising a 1st region containing a biologically active domain and another domain comprising an epitope, linked via a polypeptide to a 2nd region including a single chain antibody (SCA) having the light and heavy chains of an antibody variable region which specifically binds the epitope in the 1st region. The chimeric polypeptide assumes a conformation in which the SCA is bound to the epitope of the 1st region and protects its biological activity from deactivation by denaturing temperatures or pH conditions, proteolytic enzymes, oxidising agents or alcohol. The regions of the chimeric polypeptide interact to form a structure analogous to an antibody-antigen complex. A L-asparaginase-SCA fusion protein of the above type has better trypsin resistance than free L-asparaginase.

Sequence 1848 BP; 435 A; 498 C; 517 G; 398 T; 0 other;  
Query Match 25.9%; Score 294; DB 18; Length 1848;  
Best Local Similarity 56.3%; Pred. No. 9.7e-82; Indels 6; Gaps 1;  
Matches 574; Conservative 0; Mismatches 440;

91 GGTCCGGATCCAGCGCTCTGTTTGTATGGCTTAACCCCAAGTGCATATCTCTAGCCATC 150  
835 GGTGGTGGGTGGGTGGCGGGATCTAGATCTTACCCAATATACCACTTTTAGCAACC 894  
151 GGAGGCAACCATCGTGGTTCGGGGGAATCTAGCGTCAAGAGTAGTCTCTCTGCTGGAGCA 210  
895 GCGGGACCATTTGCGGGTGGTGGTCTCCGACATCCGACCAATCTTACTACACAGTGGTAA 954  
211 GTACCCGTTGATAAGCTTCTTCAGCGCTCCCTGCGCATCAACGACTAGCCACCATCAAG 270  
955 GTTGGCGTGAAGAAATCTGGTTAATGCGGTGCGCAACTAAAAGACATTTGCGAAGCTTAA 1014  
271 GGTGAACAGATCTCAAGCATTTGGTCCCAAGAGATGACGGTAAAGTGTGGCTTAAACTA 330  
1015 GCGGAGCAGGTAGTGAATATCGGTCTCCAGGACATGAACATATGTCTGGCTGACATG 1074  
331 GCCAAGCGTCAATGAGCTCTCTCGCCCAAAAGAGACGGAAGCCGTGTATCATCCCAT 390  
1075 GCGAAAAAATTA-----CACCGACTCGATAAGACCGGCTTCGTCAATACCCAC 1128  
391 GGAACTGACACCATGGAAGAGACGGCTTTCTTCTCAACTCACTCAAGTGAAGAGCCAAAA 450  
1129 GGTACCGACGATGGAAGAACTGCTTACTTCTCGACCTGACCGTGAATTCGACAAA 1188  
451 CCTGCTGCCCTGTAGCGGCATCGGTCCAGGCTCTTCCATAGTGTGTATGCCCATG 510  
1189 CCGGTGGTGTATGGTTCGGCGCAATGCGTCCGTCCACGTCTATGAGCGCAGACGGTCCATC 1248  
511 AATCTCTATAACCGCGTGAATGTAGCGATCAACAAAGCCTCTACTAAAGAGAGTGGT 570  
1249 AACCTGTATACCGGTAGTGACCGCAGCTGATAAAGCCTTCGCGCAACCGGTGGCTG 1308

21-JUN-1988; 88US-0205748.  
31-AUG-1992; 92US-0938505.  
23-MAY-1995; 95US-0447422.  
(HYBR-) HYBRISENS LTD.  
Ramjeesingh M, Rothstein A, Shami EY;  
WPI; 1997-558200/51.  
Self-protecting chimeric polypeptide comprising biologically active sequence and single-chain antibody sequence - has resistance to e.g. disrupting temperature, presence of proteolytic enzymes, etc.  
Example 2; Columns 27-30; 29pp; English.  
The present sequence is a chimeric gene containing the cDNA for the immunoglobulin (Ig) heavy and light chain variable regions of a murine anti-asparaginase II monoclonal antibody (MAB), and human asparaginase II. The gene was used in the preparation of a novel recombinant chimeric polypeptide, comprising a 1st region containing a biologically active domain and another domain comprising an epitope, linked via a polypeptide to a 2nd region including a single chain antibody (SCA) having the light and heavy chains of an antibody variable region which specifically binds the epitope in the 1st region. The chimeric polypeptide assumes a conformation in which the SCA is bound to the epitope of the 1st region and protects its biological activity from deactivation by denaturing temperatures or pH conditions, proteolytic enzymes, oxidising agents or alcohol. The regions of the chimeric polypeptide interact to form a structure analogous to an antibody-antigen complex. A L-asparaginase-SCA fusion protein of the above type has better trypsin resistance than free L-asparaginase.

Sequence 1848 BP; 435 A; 498 C; 517 G; 398 T; 0 other;  
Query Match 25.9%; Score 294; DB 18; Length 1848;  
Best Local Similarity 56.3%; Pred. No. 9.7e-82; Indels 6; Gaps 1;  
Matches 574; Conservative 0; Mismatches 440;

91 GGTCCGGATCCAGCGCTCTGTTTGTATGGCTTAACCCCAAGTGCATATCTCTAGCCATC 150  
835 GGTGGTGGGTGGGTGGCGGGATCTAGATCTTACCCAATATACCACTTTTAGCAACC 894  
151 GGAGGCAACCATCGTGGTTCGGGGGAATCTAGCGTCAAGAGTAGTCTCTCTGCTGGAGCA 210  
895 GCGGGACCATTTGCGGGTGGTGGTCTCCGACATCCGACCAATCTTACTACACAGTGGTAA 954  
211 GTACCCGTTGATAAGCTTCTTCAGCGCTCCCTGCGCATCAACGACTAGCCACCATCAAG 270  
955 GTTGGCGTGAAGAAATCTGGTTAATGCGGTGCGCAACTAAAAGACATTTGCGAAGCTTAA 1014  
271 GGTGAACAGATCTCAAGCATTTGGTCCCAAGAGATGACGGTAAAGTGTGGCTTAAACTA 330  
1015 GCGGAGCAGGTAGTGAATATCGGTCTCCAGGACATGAACATATGTCTGGCTGACATG 1074  
331 GCCAAGCGTCAATGAGCTCTCTCGCCCAAAAGAGACGGAAGCCGTGTATCATCCCAT 390  
1075 GCGAAAAAATTA-----CACCGACTCGATAAGACCGGCTTCGTCAATACCCAC 1128  
391 GGAACTGACACCATGGAAGAGACGGCTTTCTTCTCAACTCACTCAAGTGAAGAGCCAAAA 450  
1129 GGTACCGACGATGGAAGAACTGCTTACTTCTCGACCTGACCGTGAATTCGACAAA 1188  
451 CCTGCTGCCCTGTAGCGGCATCGGTCCAGGCTCTTCCATAGTGTGTATGCCCATG 510  
1189 CCGGTGGTGTATGGTTCGGCGCAATGCGTCCGTCCACGTCTATGAGCGCAGACGGTCCATC 1248  
511 AATCTCTATAACCGCGTGAATGTAGCGATCAACAAAGCCTCTACTAAAGAGAGTGGT 570  
1249 AACCTGTATACCGGTAGTGACCGCAGCTGATAAAGCCTTCGCGCAACCGGTGGCTG 1308

QY 571 ATTGTGATGAACGATGAGATTTCAGCGCCGAGAGAGCGACCAAGCTCAACACCCACCGCA 630  
 Db |||||  
 QY 1309 GTAGTGATGAATGACACCGTCTTGATGGCCGAGGTCAACCAACCAACACCGGAC 1368  
 Db |||||  
 QY 631 GTCAATGCAATTCGTCGCGCAACACAGAGTAAATCGGCACAGCTCTATTATGCAAGTC 690  
 Db |||||  
 QY 1369 GTAGCGACCTTCAAGTCTGTTAACCTAGCGTCTCTGGTTACATTCACAACGTAAGATT 1428  
 QY 691 GAGTATTTCACTCAATCCGTTGACCTCACCCCTGCAAGTGAGTTGATATTAGCAAA 750  
 Db |||||  
 QY 1429 GACTACACGAGTACCCGCGCATAGACATACAGCAGACAGCCCAATTCGATGCTCTAAG 1488  
 QY 751 ATCGAGAACTCCCGAGATCGATATCTTTACGCTCACCCCGATGATATGATCTTTTA 810  
 Db |||||  
 QY 1489 CTGAATGAATCCCGAAGTCCGCAATTTTATACTAGCTTAACGCAATCCGATCTTCG 1548  
 QY 811 GTCAATGAGCCCTTCAGCGAGGAGCCAAAGGAATCATCTGAGGAGGATGGCAATGGG 870  
 Db |||||  
 QY 1549 GCTAAGCACTGTAGATGCGGGCTATGTCGCTGTAGCGCTGTGGTAAACGGC 1608  
 QY 871 AACCTTTCCCTTTGACTCAAAATGCTCTTGAAGAGCAGCAATCAGGCGTGTGCTC 930  
 Db |||||  
 QY 1609 AACCTGTATAAATCTGTTCGACAGCTGGCGACCGCGGAAACCGGTACTGCGATC 1668  
 QY 931 GCTCGAAGCTCTAGAGTGGGAGTGTTCACACACCCAGAGGTGAGTGAATGATAAG 990  
 Db |||||  
 QY 1669 GTGGGTTCTCCCGCTACCGACGGGCGCTACCACTCAGGATGCCGAAGTGGATGCG 1728  
 QY 991 AACTTGGTTTGGGCTACAGAGAGTCTCAACCTCAAAAGCCAGAGTCTTCTCCACCTATTA 1050  
 Db |||||  
 QY 1729 AATACGGCTTCGTCGCTCTGGCAGCTGAACCGGCAAAAGCGCGCTTCTGCTGCAA 1788  
 QY 1051 TTAGCCCTCAACAACTAGTGTAGAGAGCGATCCAAAGATCTTCTCCACCTATTA 1110  
 Db |||||  
 QY 1789 CTGCTCTGACGCAACCAAGATCCGACGATCAGCAGATCTTCAATCAGTACTAA 1848

## RESULT 6

AA14328  
 ID AAX14328 standard; DNA; 1196 BP.

AC AAX14328;

DT 31-MAR-1999 (first entry)

DE H. pylori GHPO 187 gene.

KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease; ss.

OS Helicobacter pylori.

PH Key Location/Qualifiers  
 FT CDS 41..1135  
 FT /\*tag= a

XX WO9843478-A1.

XX 08-OCT-1998.

XX 01-APR-1998; 98WO-US06371.

XX 29-JUL-1997; 97US-0902615.

XX 01-APR-1997; 97US-0833457.

XX 24-JUN-1997; 97US-0881227.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;

XX WPI; 1998-542293/46.

XX P-PSDB; AAW98609.

XX

PT New isolated Helicobacter polynucleotides - used to develop products  
 for the diagnosis, prevention and treatment of Helicobacter  
 infections and gastrointestinal diseases

XX Claim 1; Page 1227-1229; 2054pp; English.

XX

CC This sequence represents a polynucleotide of the invention. It was  
 isolated from Helicobacter pylori and encodes a H. pylori GHPO protein.  
 The polypeptides can be used for preventing or treating Helicobacter  
 infections, and gastroduodenal diseases associated with these  
 infections, including acute, chronic, and atrophic gastritis, and peptic  
 ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used  
 for the production of antibodies. The products can also be used for  
 detection and diagnosis.

SQ Sequence 1196 BP; 355 A; 222 C; 286 G; 333 T; 0 other;

Query Match 22.9%; Score 259.2; DB 19; Length 1196;

Best Local Similarity 55.5%; Pred. No. 7.3e-71;

Matches 548; Conservative 0; Mismatches 428; Indels 12; Gaps 2;

QY 126 ACCCAAGTACTATCTCTGCGAGCAGTCCAGGAGCAGGAGTTCGGGGGATCTAGCGT 185

Db |||||

QY 157 ACCACCATTTCTTACTGGCAGACGGGGGAGATTGACAGGAGTGTGGCGGCGAG 216

Db |||||

QY 186 CAAGAGTAGTACTCTCTGCGAGCAGTCCAGGAGTGTAGTCTTTCAGCGGTCCTGC 245

Db |||||

QY 217 TTTGGGTAGTTAAGAGTGTGAGTTGGGCATCAAGAGAGCTTTGAAGGTATCCCTAG 276

Db |||||

QY 246 CATCAAGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAT 305

Db |||||

QY 277 TCTTAACAGACTCGCTCGCATTCAGGGGAGCAGATTCTAACATCGGCTCAACAGAT 336

Db |||||

QY 306 GACGGTAAAGTGTGGCTTAAACTAGCCAGGAGTGTCAATGAGCTCCTCGCCCAAAAGA 365

Db |||||

QY 337 GAATGAAGAGTATGGTTCAAGCTGCCAAACGTGCCCAAGAAATTCATAGATGATGCG 396

QY 366 GACCAAGCGGTGATCATCACCATGGAACTGACACCATGGAAGAGACGCTTCTCTCT 425

Db |||||

QY 397 TATTCAAGCGGTGTCATCAGCATGGCAGGACACTTTAGAGAGAGCGCGTATTTT 456

QY 426 CAACCTCAGCGTGAAGAGCAAAACCTGTCTGCTGTAGCGGCATGCGTCCAGGCTC 485

Db |||||

QY 457 AAACCTAGTTTACGCTCCACAAACCGGTGTGCTGGGAGCGATGCGTAATGCTGC 516

QY 486 TTCCATGAGTCTGATGCGCCCATGAATCTCTATAACGCGCTGATGTAGCGATCAACAA 545

Db |||||

QY 517 TCTTTGAGCGCGATGGGCTTTGCAATTTATATAATGCTGTAGCGTGGCTCAATGA 576

QY 546 AGCCTCTACTAACAAAGAGTGTGATGTGATCAAGATTCACGCGCCAGAGA 605

Db |||||

QY 577 AAAAAGTCCGATTAAGCGGTGTAGTGGTGATGGAGCAGTAATATTTTAGCGCTAGA 636

QY 606 AGGACCAAGCTCAACACCAACCGCAGTCAATGATTTGCTTCGCCCAACACAGCTAAAT 665

QY 637 AGTGATTAACCAACGACACCAACCGCAGTCCACCTTTAAAGCTTAATAGCGCGAT 696

QY 666 CGGCACAGTCTATTATGGCAAGTCCAGTATTTCACTCAATCCGCTTCGACTCACACCT 725

QY 697 AGGAGCGGTATATTATGGCAAAACGCGCTATTATGATGCGAGCTTTGAGAAAACACAC 756

QY 726 TGCAGTGTGATTTAGCAAAATCGAA---GAACTCCCGCAGAGTTCGATATCTTTTA 782

QY 757 AGAGAGCGAATTTCCCTTTCAACTCAAAACCCCTGCTTAAAGTGATATTTTA 816

QY 783 CGCTACCCCGATGATGATGTTTGTAGTCAATGAGCGCTTCAGGAGGAGCGCAAGG 842

QY 817 CACGATGCTGCATGACCCCTGATTTATTCGAAGCGAGCTTAACTCGCATGCAAAAG 876

QY 843 AATCATCCATGAGGAGTGGCAATGGGAAACCTTTTCCCTTTGACTCAAAATGCTTTGA 902

QY 877. CGTTGTGATAGCCGGGTGGGTAAATGGGAATGTGAGCGCTGGGTTTTTAAAGCGATGCA 936



Db 1681 ATTCAAGAGTATTTCATATTCAT 1708

RESULT 8

AAQ68438  
ID AAQ68438 standard; DNA; 1014 BP.

XX AAQ68438;

XX 25-MAR-2003 (updated)

DT 12-JAN-1995 (first entry)

XX Pseudomonas glutaminase gene.

DE Glutaminase; antiviral; virucide; anticancer; cancer therapy;  
KW HIV virus; gene therapy; Escherichia coli; ds.

XX Pseudomonas sp.

XX WO9413817-A1.

XX 23-JUN-1994.

PF 04-DEC-1992; 92WO-US10421.

XX 04-DEC-1992; 92WO-US10421.

PR 04-DEC-1992; 92AU-0032358.

XX (MEME-) ME MEDICAL ENZYMES AG.

XX Freeman AG, Macallister TW, Roberts J, Sethuraman N;

XX WPI; 1994-217891/26.

XX P-PSDB; AARS9739.

Recombinant glutaminase derived from Pseudomonas 7A - expressed  
in E. coli to increase yield and avoid Pseudomonas endotoxins for  
antiviral and anticancer therapy

PS Disclosure; Page 33; 60pp; English.

XX Chromosomal DNA from Pseudomonas sp. 7A (ATCC 29598) was used to  
construct a genomic library in Escherichia coli LE392. Screening  
with mixed oligonucleotide probes was used to isolate a glutaminase-  
encoding clone. This was sequenced using the primers given in  
AAQ68439-47. The gene can be used to manufacture recombinant  
glutaminase, free of Pseudomonas exotoxin, for use in e.g. HIV and  
cancer therapy. The gene may also be used in gene therapy protocols.  
(Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 1014 BP; 226 A; 324 C; 318 G; 146 T; 0 other;

Query Match 19.0%; Score 215.8; DB 15; Length 1014;

Best Local Similarity 52.6%; Pred. No. 3.3e-57;

Matches 519; Conservative 0; Mismatches 462; Indels 6; Gaps 2;

QY 128 CCCAAGTGAATCTTCCTAGGACAGGAGGACCATCTGCTTGGGGGAATCTAGCGTCA 187

Db 29 CCAAGTGTGTATCTTGGCCACCGCGGACCATCTGCGCGCTGCGCGCGCGCA 88

QY 188 AGAG---TAGCTACTCTGTGGAGCAGTCACTGCTTGTATAGCTTCTTGCAGCGCTCCCTG 244

Db 89 ACAGCGCCACCTACCAAGGCTGCCAAGTTTGGGTTCGACAGCTGATGCGGGTGGCCG 148

QY 245 CCATCAACCACTAGCCACCATCAGGTTGACAGATCTCAAGCATTTGGCTCCCAAGAGA 304

Db 149 AGCTGGCCACCTTGGCCATGTCGCGCGGACGAGTGTGATGCGCTCCGAAAGCA 208

QY 305 TGACGGGTAAAGTGTGGCTTAACTAGCAACGCTGTCAATCAGCTCTCTCGCCAAAG 364

Db 209 TCACCAACGACGACCTGTCTAGCTGGACAGGAGTGTGCGCGAGTGTGCGCGACGAAATG 268

QY 365 AGACGGAAGCGGTGATCATATCCCATGGAAGTGCACCATGTGAAGACCGCTTCTTCC 424

Db 269 ACGTCGATGGCATCGTCATCCCTTGGCACGACACCTTGGAGAAACGCGCTATTTT 328

QY 425 TCAACCTCAGCGTGAAGACCAAAACCTGTCTGCTTGTAGCGCCATGCGTCCAGGCT 484

Db 329 TGAACCTCTGTGAAAGACCGACAAAGCGGATCGTGTGCTGCTTCATGCGCGCGCA 388

QY 485 CTTCCATGAGTCTGATGGCCCCCATGAATCTCTATAACGCGCTGAATGTAGGATCAACA 544

Db 389 CGCCATGTTCGCCGACGCGCATGCTCAACCTGTACACGCGCTGGCGGTGGCCAGCA 448

QY 545 AAGCCTCTACTTAAACAGAGTGTGATGTGAACAGATGAGATTTCAGCGCCAGAG 604

Db 449 AGGACTCGCGCGCAAGGCGCTGTGTCACCATGAACGACGAGATCCAGTCCGCGCGTG 508

QY 605 AAGCGACCAAGCTCAACACCAACCGCAGTCAATGATTTGCTTCGCGCCACACAGTAAA 664

Db 509 ACGTGAGCAAGTTCGATCAACATCAAGACCGAAGCCTTC---AAGAGCGCTTGGGCGCGC 565

QY 665 TCGGCACAGTCTATTATGGCAAGTTCGAGTATTTCACTCAATTCGCTTCCAGCTCACACC 724

Db 566 TGGCATGTGTGTGAAGGCAAGTCTGCTGCTTCCGCTTCCGCGCCCAAGCGCACACGG 625

QY 725 TTGCAAGTGTGATTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTTCGATTTTTCG 784

Db 626 TCAACTCGAGTTCGACATCAAGCAGATCAGACGCTGCGCCAGTGTGACATCGCTACA 685

QY 785 CTCACCCGATGATCTGATGTTTGTAGTCAATGAGCGCTTTCAGCGAGGAGCCAAAGAA 844

Db 686 GCTATGCAACGCTCACCGACACCGCTTCAAGGCGCTTGGACAGAAACGCGCGCAAGGCGC 745

QY 845 TCATCCATGAGGATGGGCAATGGGAAACCTTTCCCTTTTGTACTCAAAATGCTCTTGA 904

Db 746 TGATCCATGCGCGCACCGGCAATGGCTCGCTGCTGCGGGTGTGCGCGCTGCGCAGCCTG 805

QY 905 AAGCAGCAAAATCAGGCGTGTGCTGCTCGAAGCTCTAGAGTGGGAGTGTTCACCA 964

Db 806 AGCTGCGCAAGAACGCGGTGCAGATCATTCGTCGTCAACAGGGCGGTTTCGTGC 865

QY 965 CCCAAGGCTGAAGTGGATGATAAGAACTTGGTTTGTGGCTTACAGAGATGTCAACC 1024

Db 866 TGGGTAAACCGGAGCAGCGCGACAGAAACGACTGGGTGCTGGCGCCACGACTGAAAC 925

QY 1025 CTCAAAAGCCAGAGTCTTCTTATGTAGCTTACCCCTACCAAACTAGTGTAGAGAGCGA 1084

Db 926 CGCAGAGGCGCCATCTTGGCGATGGTGGCAATGACCAAGACCCAGGACAGCAAGGAGC 985

QY 1085 TCAAAAGATCTTCTCCACCTATTAAAT 1111

Db 986 TGCAGCGCATTTTCTGGGAATACTGAT 1012

RESULT 9

ABZ80801

ID ABZ80801 standard; DNA; 1174 BP.

XX ABZ80801;

XX 30-MAY-2003 (first entry)

XX Erwinia carotovora L-asparaginase gene.

XX gene; ds; asparaginase; cytostatic; blood; acute lymphoblastic leukemia.

XX Erwinia carotovora.

XX Key Location/Qualifiers

XX CDS 48..1094

XX /\*tag= a

XX /product= "L-asparaginase"

XX /EC number= 3.5.1.1

XX primer\_bind 48..72

XX /\*tag= b



CC Helicobacter pylori selected from 421 complexes given in the  
 CC specification. The complex of protein-protein interactions are useful  
 CC for screening for agents which modulate the interaction of proteins.  
 CC Modulating compounds which binds to a targeted bacterial protein may be  
 CC used for treating or preventing ulcers in a human or animal. This  
 CC sequence encodes a selected interacting domain (SID), identified via  
 CC protein-protein interactions.

XX SQ Sequence 345 BP; 90 A; 69 C; 104 G; 82 T; 0 other;

Query Match 9.6%; Score 108.8; DB 24; Length 345;  
 Best Local Similarity 57.7%; Pred. No. 1.1e-23;  
 Matches 194; Conservative 0; Mismatches 142; Indels 0; Gaps 0;  
 QY 142 CTAGCCACAGGAGGACCATCGCTGGTTCGGGGGATCTAGCGTCAAGAGTAGCTACTCT 201  
 DB 2 CTGGCGACAGGGGACGATTCAGAGGAGTGTGCGAGCGGAGTTGGTGTATTAAG 61  
 QY 202 GCTGGAGGAGTCACCGTTGATTAAGCTTCTTGCGCGCTCCCTGCCATCAACGACCTAGCC 261  
 DB 62 AGTGTGAGTTGGGCATCAAGAGCTTTTGAAGGCTATCCCTAGTCTTAAAGACTCGCT 121  
 QY 262 ACCATCAAGGTTGAACAGATCTCAAGCATTTGCTCCCAAGAGATGACGGTAAAGTGTGG 321  
 DB 122 CGAATTCAGGGGACGATTTCTAACATCGGCTCACAGACATGAATGAAGAGGTATGG 181  
 QY 322 CTTAAACTAGCAAGCGGTGTCAATGAGCTCTCGGCCCAAAAGAGACCGCGGTATC 381  
 DB 182 TTCAAGCTCGCCAAAGTGGCCCAAGATTGCTAGATGATAGCGGTATTCAGCGGTGTC 241  
 QY 382 ATCAACCATGAATGACATGACATGAGAGACCGCTTCTTCTCAACTCAGCGTGAAA 441  
 DB 242 ATCAAGCATGCGACGACATTTTGAAGAGAGCGGTATTTTAAACTTAGTTTATAGC 301  
 QY 442 AGCCAAAACCTGTCTGCTTGTAGCGCCATGCGT 477  
 DB 302 TCCAAAACCGTGTGCTGTGGAGCGATGCGT 337

RESULT 11

AAS21124

ID AAS21124 standard; cDNA; 543 BP.

AC AAS21124;

XX 21-MAY-2002 (first entry)

DE CDNA encoding Tm 3.9/His tag fusion protein minus signal peptide.

XX Anti-freeze peptide; Tm 3.9; yellow mealworm beetle; THP; cryoinjury;  
 KW thermal hysteresis protein; cryosurgery; cold climatization; ss;  
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;  
 KW transgenic animal; His tag.  
 XX Tenebrio molitor.

Key Location/Qualifiers  
 CDS 64..513  
 FT /\*tag= a  
 FT /product= "Tm 3.9 clone protein"  
 FT misc\_feature 64..165  
 FT /\*tag= b  
 FT /note= "Nucleotides encoding His tag"  
 FT mat\_peptide 166..510  
 FT /\*tag= c

XX WO200194378-A1.

PN 13-DEC-2001.

XX 07-JUN-2001; 2001WO-US18532.

XX 08-JUN-2000; 2000US-210446P.

XX PR

XX (UNYV ) UNIV NEW YORK STATE RES FOUND.  
 PA (HORN/) HORWATH K L.  
 PA (MYER/) MYERS K L.  
 PA (EAST/) EASTON C M.

XX Horwath KL, Myers KL, Easton CM;  
 PI WPI; 2002-090137/12.  
 DR P-PSDB; AAU10053.

XX New CDNA polynucleotide encoding a thermal hysteresis protein which is  
 PT a Type III anti-freeze protein derived from the Tenebrionidea  
 PT Superfamily, useful for providing antifreeze protection to improve the  
 PT quality of food -

XX Claim 4; Page 345; 364pp; English.

XX This invention relates to a CDNA polynucleotide comprising a nucleotide  
 CC sequence encoding a thermal hysteresis protein (THP) which is a Type III  
 CC anti-freeze protein derived from the Tenebrionidea Superfamily. Anti-  
 CC freeze proteins lower the freezing point of a solution without affecting  
 CC the melting point of the solution. An activated anti-freeze protein may  
 CC be incorporated into plant, produce or fish in an amount sufficient to  
 CC provide antifreeze protection or in a region of a target tissue to  
 CC provide antifreeze protein to limit tumour cell or target tissue to  
 CC cryoinjury during cryosurgery. The proteins of the invention may also be  
 CC used in hypothermic solutions or bathing media to reduce cold damage in  
 CC order to provide cryogenic or hypothermic preservation of cells and  
 CC tissues. The proteins may be used as de-icing formulations or used on  
 CC surfaces to reduce existing ice buildup or abate the formation of ice  
 CC buildup on surfaces such as a road, aircraft, household products,  
 CC machinery and plant surfaces or as a food product to improve the quality  
 CC of food by abating freezing of solutions, freezer burn, or degradation  
 CC due to cold storage. The polynucleotides for the activated protein can  
 CC be used to create transgenic or gene-modified plants, crops, fish, or  
 CC animals having greater tolerance to cold climatization. The Tm 12.86  
 CC antibody/antiserum which is also used as a screening device to screen  
 CC cDNA libraries in an expression system, including cross-species cDNA  
 CC libraries to identify homologous sequences in other species. The  
 CC present sequence represents a CDNA encoding the Tm 3.9 clone/His  
 CC tag fusion protein minus the signal peptide of the invention. This CDNA  
 CC was created to facilitate purification and to try to enhance the anti-  
 CC freeze activity of the recombinant protein.

XX Sequence 543 BP; 164 A; 133 C; 144 G; 102 T; 0 other;

Query Match 9.1%; Score 103.6; DB 24; Length 543;  
 Best Local Similarity 88.9%; Pred. No. 5.9e-22;  
 Matches 112; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATGGGACAGCGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60  
 DB 64 ATGGGACAGCGCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 123

QY 61 ATGGCTAGCATGACTGGTGGACAGCAAAATGGGTGGCGGATCCAGCGCCCTCTGTTTCATG 120  
 DB 124 ATGGCTAGCATGACTGGTGGACAGCAAAATGGGTGGCGGATCCCTACCGATGAACAGATA 183

QY 121 GCTAAA 126  
 DB 184 CAGAAA 189

RESULT 12

AAS21116

ID AAS21116 standard; cDNA; 543 BP.

XX AAS21116;

XX 21-MAY-2002 (first entry)

XX CDNA encoding Tm 2.2 clone/His tag fusion protein minus signal sequence.









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:41:09 ; Search time 105,214 Seconds  
(without alignments)  
4753,042 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133

Sequence: 1 atggcagcagccatcatca.....aagaaagggaattctttcac 1133

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfileseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1129.8	99.7	1133	3	US-09-094-435-3 Sequence 3, Appli
2	296.4	26.2	1830121	4	US-09-557-884-1 Sequence 1, Appli
3	296.4	26.2	1830121	4	US-09-643-990A-1 Sequence 1, Appli
4	294	25.9	1848	1	US-08-447-422-15 Sequence 15, Appli
5	215.8	19.0	1014	5	PCT-US92-10421-1 Sequence 1, Appli
6	210.8	18.6	1017	4	US-08-050-482A-1 Sequence 1, Appli
7	200.4	17.7	1182	4	US-09-252-991A-5793 Sequence 5793, Ap
8	200.4	17.7	11335	4	US-09-252-991A-5877 Sequence 5877, Ap
9	159.2	14.1	1125	4	US-09-328-352-2336 Sequence 2336, Ap
10	147.4	13.0	2748	4	US-09-252-991A-5773 Sequence 5773, Ap
11	101.4	8.9	552	4	US-09-230-078A-3 Sequence 3, Appli
12	80	7.1	270	4	US-09-252-991A-5837 Sequence 5837, Ap
13	76.6	6.8	1005	4	US-09-107-532A-1293 Sequence 1293, Ap
14	74	6.5	2231	4	US-08-961-527-265 Sequence 265, Appl
15	72.6	6.4	366	4	US-09-613-303-11 Sequence 11, Appl
16	71.4	6.3	981	4	US-09-134-001C-397 Sequence 397, Appl
17	70.4	6.2	1230	4	US-09-613-303-54 Sequence 54, Appl
18	66.2	5.8	2847	4	US-09-613-303-20 Sequence 20, Appl
19	65.2	5.8	1989	4	US-09-207-388-4 Sequence 4, Appli
20	64.4	5.7	2016	3	US-09-198-723A-110 Sequence 110, Appl
21	64.4	5.7	2016	3	US-09-198-723A-111 Sequence 111, Appl
22	63.8	5.6	1624	4	US-09-647-224A-13 Sequence 13, Appl
23	63.8	5.6	1707	4	US-09-207-388-10 Sequence 10, Appl
24	62.8	5.5	648	3	US-09-198-723A-112 Sequence 112, Appl
25	62.8	5.5	648	3	US-09-198-723A-113 Sequence 113, Appl
26	62.8	5.5	648	3	US-09-198-723A-114 Sequence 114, Appl
27	62.8	5.5	648	3	US-09-198-723A-115 Sequence 115, Appl

Sequence 92, Appl  
Sequence 93, Appl  
Sequence 94, Appl  
Sequence 95, Appl  
Sequence 96, Appl  
Sequence 97, Appl  
Sequence 98, Appl  
Sequence 99, Appl  
Sequence 100, Appl  
Sequence 101, Appl  
Sequence 102, Appl  
Sequence 103, Appl  
Sequence 104, Appl  
Sequence 105, Appl  
Sequence 106, Appl  
Sequence 107, Appl  
Sequence 108, Appl  
Sequence 109, Appl

28 62.8 5.5 651 3 US-09-198-723A-92  
29 62.8 5.5 651 3 US-09-198-723A-93  
30 62.8 5.5 651 3 US-09-198-723A-94  
31 62.8 5.5 651 3 US-09-198-723A-95  
32 62.8 5.5 651 3 US-09-198-723A-96  
33 62.8 5.5 651 3 US-09-198-723A-97  
34 62.8 5.5 651 3 US-09-198-723A-98  
35 62.8 5.5 651 3 US-09-198-723A-99  
36 62.8 5.5 651 3 US-09-198-723A-100  
37 62.8 5.5 651 3 US-09-198-723A-101  
38 62.8 5.5 1998 3 US-09-198-723A-102  
39 62.8 5.5 1998 3 US-09-198-723A-103  
40 62.8 5.5 1998 3 US-09-198-723A-104  
41 62.8 5.5 1998 3 US-09-198-723A-105  
42 62.8 5.5 1998 3 US-09-198-723A-106  
43 62.8 5.5 1998 3 US-09-198-723A-107  
44 62.8 5.5 1998 3 US-09-198-723A-108  
45 62.8 5.5 1998 3 US-09-198-723A-109

#### ALIGNMENTS

#### RESULT 1

US-09-094-435-3  
; Sequence 3, Application US/09094435  
; Patent No. 6251388  
; GENERAL INFORMATION:  
; APPLICANT: Donald L Durden  
; TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES  
; TITLE OF INVENTION: ASPARAGINASE IN THE TREATMENT OF HUMAN  
; TITLE OF INVENTION: HEMATOLOGIC AND AUTOIMMUNE DISEASE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq for Windows 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/094,435  
; FILING DATE: Filed Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/049,085  
; FILING DATE: June 9, 1997  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 234/274  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1133 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-094-435-3

Query Match 99.7% Score 1129.8; DB 3; Length 1133;

[illegible]

```

Db      1021 AACCCCTCAAAAAGCCAGAGTGCTTCTTATGTTAGCCCTCACAAAACCTAGTATAGAGAG 1080
QY      1081 GCGATCCAAAAGATCTTCTCCACCTATTAAATCAAGAAAGGGAATCTCTTCCAC 1133
        |||||||
Db      1081 GCGATCCAAAAGATCTTCTCCACCTATTAAATCCAGAAAGGGAATCTCTTCCAC 1133
        |||||||

RESULT 2
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Haemolysate sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

```

	Query Match	26.2%;	Score 236.4;	DB 4;	Length 1830121;	
	Best Local Similarity	57.1%;	Pred. No. 2.9e-86;			
	Matches 562;	Conservative	0;	Mismatches 417;	Indels	Gaps 1;
QY	126	ACCCAAAGTGA	TACTCTCTAGCCACAGAGGACCA	TCTCGTTCGGGGGAATCTAGCGT	185	
Db	802722	ACCAATATTACA	ATCTTGGCAACGGGTGTACCA	TTCGAGGAAGCGGCAAGTTCGGT	802788	
QY	186	CAGAGTAGCT	TCTCTGTGGAGCAGTCCCGTTGAT	AAGCTTCTTGCAGCCGTCCTCGTGC	245	
Db	802782	AAATCTCGCGT	ATAAAGTGGACAAATTAAGTATTG	ATACTTTAATTAAGCTGTACCA	802844	
QY	246	CATCAACGCT	AGCCACCACTCAAGGTGAACAGAT	CTCAAGCATTTGGCTCCCAAGAGAT	305	
Db	802842	AATCAAAAAT	TGCCCAACATTAAGGTGACAAAT	TGTAATAATAGTTTCAAGACAT	802908	
QY	306	GACGGTAA	GGTGTGGCTTTAACTAGCAACGGT	GTCAATGAGTCTCTCGCCCCAAAAAGA	365	
Db	802902	GAATGACGA	AGTCTTGGCTTAAACTTGGCAAAAG	CCCATCAATGTCT-----AATGTAAAAAG	802955	
QY	366	GACCGAAG	CCGCTGATCATCA	CCCCATGCAACTGACACCATGAAGAGACCGCTTTCTTCTCT	425	

Thu Aug 28 09:00:59 2003

us-09-937-982-3.rni

802956 TACTGATGGATTGTCATTACCCATGTCAGATACCATGGAAGAGACGGCTATTTCCT 803015  
426 CAACCTCAGCGTGAAGAGCAAAACCTGCTGCTGTAGGCGCAATGCGTCCAGGCTC 485  
803016 AGATTTAACCGTAAATGTGAAGAACCGGTGTTCTGTTGGGCAATGCGTCTGCAAC 803075  
486 TTCATGAGTGTGATGCGCCCTGATCTCTATAACCGCGTGAATAGGATCAACAA 545  
803076 AGAAAGAGTGTGATGCGCCATTAATCTTTACATGCTGTGCTGCGCAGACAA 803135  
546 AGCCTCTACTAAACAAAGAGTGTGATGATGAACGATGAGATTACCGCGCAGAGA 605  
803136 AAAATCAAGTGTGCTGCTGTTTGTAGTCGCAATGAATTAAGAGTCTAGTGTCCGA 803195  
606 AGGACCAAGCTCAACACCGCGAGTCAATGCAATTTGCTTCGCCCAACACAGTAAAT 665  
803196 TGTAAACAAACAGTAGACCGCAGTGCACACGTTCCATTCACCAATATGTTCTCT 803255  
666 CGGCACAGTCTATTATGCAAGTGTGATGATGATGATGATGATGATGATGATGAT 725  
803256 AGGCTATATTATACACGAAAGTGGACTATGAACGTTCCCGAAGAACACATACAT 803315  
726 TGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785  
803316 CAACACTCCGTTTAAACGATGAGAAATTTAGATAGCTACCCAAAGTGGGATTTATGC 803375  
786 TCACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845  
803376 TTAATCAATGCACTGTGCAACCAATTAAACGCAATTAATCAATGCTGCTATCAAGGAT 803435  
846 CATCATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 905  
803436 TGTATCTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 803495  
906 AGGACCAATCAGCGTGTGCTGCTGCAAGCTCTAGAGTGGGAGTGGTTCACAC 965  
803496 AGCCCAAGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1025  
966 CCAAGAGGCTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085  
803556 TCGTGAAGCTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803615  
1026 TCAAAAGCCAGTGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085  
803616 ACAGAAAGCCGCTGCTTGCATTTAGCTTAACTCAACTCAAGATCCNAAAGTAA 803675  
1086 CCAAAAGATCTTCTCCACCTATTAA 1110  
803676 TCAACATATTTGAGAGATTTCTAA 803700

RESULT 3  
US-09-643-990A-1  
; Sequence 1, Application US/09643990A  
; Patent No. 6528289  
; GENERAL INFORMATION:  
; APPLICANT: Robert D. Fleischmann  
; Mark D. Adams  
; Owen White  
; Hamilton O. Smith  
; J. Craig Venter  
; TITLE OF INVENTION: The Nucleotide sequence of  
; the Haemophilus influenzae Rd Genome, Fragments  
; thereof, and Uses Thereof  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville,  
; STATE: MD  
; COUNTRY: USA  
; ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/643,990A  
FILING DATE: 23-Aug-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,429  
FILING DATE: 1995-06-07  
APPLICATION NUMBER: 08/426,787  
FILING DATE: 1995-04-21  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PE186P1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-610-5790  
TELEFAX: 310-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-643-990A-1

Query Match 26.2%; Score 296.4; DB 4; Length 1830121;  
Best Local Similarity 57.1%; Pred. No. 2.9e-86;  
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;  
QY 126 ACCCCAGTGTACTCTAGCCAGGAGGACCAATCGCTGTTCCGGGGAAATCTAGCCT 185  
DB 802722 ACCAAATTAACATCTTGGCAACCGGTGGTACCATTCGAGGAGCGGCAAGTTCGT 802781  
QY 186 CAAGAGTAGTACTCTCTGGAGAGTCAACCGTGTGATGATGATGATGATGATGATGATGAT 245  
DB 802782 AAATTCGCTGATTAAGCTGGACAAATTAAGTATTGATGATGATGATGATGATGATGATGAT 802841  
QY 246 CATCAACGACCTAGCCACCAATCAAGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGAT 305  
DB 802842 AATGAAAATATGCGCAATTAAGAGTGAAGCAATTTGAAAATAGTTTCAAGACAT 802901  
QY 306 GACGGTGAAGTGTGGCTTTAAACTAGCCAAAGCGTGTCAATGAGTCTCTCGCCCAAAAGA 365  
DB 802902 GAATGACGAGTCTGGCTAAACTGGCAAAAGCCATCAATGCTC-----AATGTAAG 802955  
QY 366 GACCGAAGCGGTGATCATCACCATGGAAGTCAACCATGGAAGTGAAGAGACCGCTTCTTCT 425  
DB 802956 TACTGATGATTTGTCTATTCATCCATGCTGATGATGATGATGATGATGATGATGATGATGAT 803015  
QY 426 CAACCTCAGGTGAAGAGCAAAACCTCTGCTGCTTGTAGGCGCATGCGTCCAGGCTC 485  
DB 803016 AGATTTAACCGTAAATGTGAAGAACCGGTGTTCTGTTGGGCAATGCGTCTGCAAC 803075  
QY 486 TTCCATGAGTGTGATGCGCCCTGATGATCTCTATAACCGCGTGAATAGGATCAACAA 545  
DB 803076 AGAAAGAGTGTGATGCGCCATTAATCTTTACATGCTGTGCTGCGCAGACAA 803135  
QY 546 AGCCTCTACTAAACAAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 605  
DB 803136 AAAATCAAGTGTGCTGCTGTTTGTAGTCGCAATGAATTAAGAGTCTAGTGTCCGA 803195  
QY 606 AGGACCAAGCTCAACACCGCGAGTCAATGCAATTTGCTTCGCCCAACACAGTAAAT 665  
DB 803196 TGTAAACAAACAGTAGACCGCAGTGCACACGTTCCATTCACCAATATGTTCTCT 803255  
QY 666 CGGCACAGTCTATTATGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 725  
DB 803256 AGGCTATATTATACACGAAAGTGGACTATGAACGTTCCCGAAGAACACATACAT 803315

QY 726 TGAAGTGAAGTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTCGATATCTTTACGC 785  
Db 803316 CAACACTCCGTTTAAACGTAGAAAATTTAGATAGCTACCCAAAGTGGGATTTATTCG 803375  
QY 786 TCACCCCGATGATCTAGTGTCTTTAGTCAATGAGCGCTTCAGCGAGGAGCCAAAGGAAT 845  
Db 803376 TTATTCAATGCACTGTGCAACATTTAAACGCAATTAATCAATGCTGGCTATCAAGGAT 803435  
QY 846 CAPCCATGAGGATGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTTTGAAA 905  
Db 803436 TGTATCTGCAGAGTTGGCAATGGAATGTTAATGCTGCACATTTAGATCGCTTAGAAA 803495  
QY 906 ACAGAGCAATCAGCGGTAGTCTGCTCGAAGCTCTAGAGTGGCGAGTGTTCACAC 965  
Db 803496 AGCCGCAAAAGATAGCGCTGCTTGTAGTCCGTTCTCTGCTGTACCAACGGTTATACA 803555  
QY 966 CAAAGAGGCTGAAGTGGATGATAAGAACTTGGTTTTGTTGGCTACAGAGTCTCAACC 1025  
Db 803556 TCGTAGCGTGAAGTTGATAGTAAATATGGCTTTGTAGCATCAGGTACTTTAAATCC 803615  
QY 1026 TCAAAAGCCAGAGTCTCTTATGTTAGCCCTCACCAAACTAGTATAGAGAGCGAT 1085  
Db 803616 ACAGAAAGCCCGTCTCTTCAATTAAGTTAACTCAAACTAAGATCCMAAAGTAA 803675  
QY 1086 CCAAGAGATCTTCTCCACCTATTA 1110  
Db 803676 TCAACAATATTTGGAAGACTTCTAA 803700

## RESULT 4

US-08-447-422-15  
; Sequence 15, Application US/08447422  
; Patent No. 5686579  
; GENERAL INFORMATION:  
; APPLICANT: SHAMI, Ezekiel Y.  
; APPLICANT: ROTHSTEIN, Asar  
; APPLICANT: RAMJESINGH, Mohabir  
; TITLE OF INVENTION: Use of Antibody/Antigen Interactions To  
; TITLE OF INVENTION: Protect or Modulate Biological Activity  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/447,422  
; FILING DATE:  
; FILING DATE: 31-AUG-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/081,410  
; FILING DATE: 22-JUN-1993  
; APPLICATION NUMBER: US 07/938,505  
; FILING DATE: 31-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/205,748  
; FILING DATE: 21-JUN-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17923/102 HYLI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1848 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-447-422-15  
  
Query Match 25.9%; Score 294; DB 1; Length 1848;  
Best Local Similarity 56.3%; Pred. No. 2.6e-87;  
Matches 574; Conservative 0; Mismatches 440; Indels 6; Gaps 1;  
  
QY 91 GGTGCGGATCCAGCGCTCTCTTTTGTAGGTTAAACCCCAAGTGAATCTCTAGCACA 150  
Db 835 GGTGTTGGTTCGGGTGGCGGATCTAGATCTTACCAATATACCAATTTTAGCAAC 894  
QY 151 GGAGGACCATCGCTGGTTCGGGGGAATCTAGCGTCAAGAGTAGTCTCTCTGTGAGCA 210  
Db 895 GCGGGGACCATTTGCGGTGGTGACTCCGCAACCAATCTACTACAGTGGGTAA 954  
QY 211 GTCACCGTTGATAAGCTTCTTGCAGCGTCCCTGCCATCAACGACCTAGCCACCATCAAG 270  
Db 955 GTTGGCGTAGAAAATCTGTTAAATGCGGTGCGCAACTAAAGACATTCGCAACGTAA 1014  
QY 271 GGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATGACGGTTAAGGTGGCTTAACTA 330  
Db 1015 GCGAGCAGGTAGTGAATATCGGTCCCGAGCATGAACGATAATGTCTGGCTGACATG 1074  
QY 331 GCCAAGCGTGAATGAGTCTCTCGCCCAAAAGAGACCGGCGTGTATCATCACCCAT 390  
Db 1075 GCGAAAAAATAA-----CACCGACTGCGATAAGACCGCGCTTCTGTCATACCCAC 1128  
QY 391 GGAATGACACATGGAAGACCGCTTCTTCTCAACCTCACGTTGAAAAGCCAAA 450  
Db 1129 GGTACCGACAGTGAAGAAATCTGTTACTTCTCGACCTCGCGTGAATTCGACAAA 1188  
QY 451 CCTGTCTGCTTGTAGCGCCATCGCTCCAGCTCTTCCATAGTGTGTGAGGCCCCATG 510  
Db 1189 CCGGTGTGTGTTGTCGGGCAATCGTCCGTCACGTTCTATGAGCGGACAGCGTCCATTC 1248  
QY 511 AATCTCTAAGCGCGTGAATGTAGCATCAACAAAGCTCTACTAAAGAGTGTGTG 570  
Db 1249 AACCTGTATAACCGGTAGTGAACGAGCTGATAAGCTTCGCAACCGTGGGTGTG 1308  
QY 571 ATTGTGATGAACATGAGATTTCACGCGCGCAGAGAGCGACAGCTCAACACACCGCA 630  
Db 1309 GTAGTGAATGACACCGTGTGTGAGCGGTGACGTCAACCAACCAACACCGCAC 1368  
QY 631 GTCAATGATTTGCTTGGCCCAACACAGGTAAATCGGCACAGTCTATTATGCAAGATC 690  
Db 1369 GTAGCGACCTTCAAGTCTGTAACTAGCGTCTCTGGTTTACATTCACACGGTAAAT 1428  
QY 691 GAGTATTTCACTCAATCCGTTTCGACCTCACACCTTGCAGTGTGATTTGATTAAGCAA 750  
Db 1429 GACTACAGGTACCCCGCACGTAAGCATACCGACGACACGCTTCTCTTAAG 1488  
QY 751 ATCGAAGAACTCCCGAGAGTCGATATTCTTTACGCTCACCCGATGATATCTGATTTTA 810  
Db 1489 CTGAATGAATGCGGAAAGTTCGCAATTTTAACTACGTTAAGCATCCGATCTTCG 1548  
QY 811 GTCAATGACGCTTCAGGAGAGGAGCAAGGAATCATCAATGACGAGTGGCAATGGG 870  
Db 1549 GCTAAAGCACTGGTAGATCGGGCTATGATGCGCATCGTTAGCGCTGGTGGGTAAAG 1608  
QY 871 AACCCCTTTCCCTTTGACTCAAAATGCTTTGAAAAAGCAGCAAAATCAGCGTAGTCGTC 930  
Db 1609 AACCTGTATAATCTGTGTTGCACGCTGGGACCGCCGCAAAACCGGTACTGAGTC 1668  
QY 931 GCTCGAAGCTCTAGAGTGGGAGTGGTTCCACACCCCAAGAGGCTGAAGTGAATGAAG 990  
Db 1669 GTGCGTTCTTCCCGGTACCGACGCGGCGTACCACTCAGGATGCCGAAGTGGATGCG 1728  
QY 991 AAACTTGGTTTGTGCTACAGAGTCTCAACCTCAAAAGCCAGAGTGTCTTCTATG 1050

Thu Aug 28 09:00:59 2003

Db 1729 AATACGGCTTCGTGCTCTGCGCTGCGCTGAAACCGCAAAAGCGCGCTTCTGCTGCAA 1788  
Qy 1051 TTAGCCCTCACAAACTAGTATGATAGAGCGCGATCCAAAGATCTTCTCCACCTATTAA 1110  
Db 1789 CTGGCTCTGACGCAACCAAGATCCGAGCAGATCTTCAATCAGTACTAA 1848

RESULT 5

PC1-US92-10421-1  
; Sequence 1, Application PC/TUS9210421  
; GENERAL INFORMATION:  
; APPLICANT: Roberts, Joseph  
; APPLICANT: MacAllister, Thomas W.  
; APPLICANT: Sethuraman, Natarajan  
; APPLICANT: Freeman, Abbie G.  
; TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND  
; TITLE OF INVENTION: ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner, Birch, McKie and Beckett  
; STREET: 1001 G Street N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20001

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/10421  
; FILING DATE: 19921204  
; CLASSIFICATION:

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P4140003.8  
; FILING DATE: 04-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 00100.41200  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1014 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: PSEUDOMONAS 7A  
; PC1-US92-10421-1

Query Match 19.0%; Score 215.8; DB 5; Length 1014;  
Best Local Similarity 52.8%; Pred. No. 1.9e-61;  
Matches 519; Conservative 0; Mismatches 462; Indels 6; Gaps 2;  
Qy 128 CCCAAGTGACTATCTAGCCACAGGAGCACCATCGCTGGTTCGGGGGAATCTAGGTCA 187  
Db 29 CCAAGTGGTGAATCTGGCCACCGCGGCAACCATCGCGCGCTGCGCGCCAGCGGGCA 88  
Qy 188 AGAG---TACTACTCTGTGGAGAGTACCGCTTGTATAGCTTCTTGAGCGGCTCCCTG 244  
Db 89 ACAGCGCCACTACAGGCTGCCAAGGTTGGCGTGCACAGCTGATGCGGGGCTCCCG 148  
Qy 245 CCATACAGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGA 304  
Db 149 AGCTGGCGGACCTGGCCCAATGTGGCGGCGAGGATGATGCGATCGCTCCGAAAGCA 208  
Qy 305 TGACGGGTAAGGTGTGGCTTAAACTAGCCAAAGCGTGTCAATGAGCTCTCGCCCAAAG 364

Db 209 TCACCAAGCAGACCTGCTCAAGTGCAGAGCGGTGGCCGAGCTGGCCGACAGCAATG 268  
Qy 365 AGACCGAAGCCGTGATCATCCCATGGAATCAGACCACTGGAAGAGACCGCTTCTTCTTC 424  
Db 269 ACGTCGATGCGTCTCATCCCATGGCACCACACCTTGGAAAGAACCGCTTCTTCTT 328  
Qy 425 TCACCTCAGCGTGAAGGCAAAACCTCTGCTGCTTGTAGGCGCCATCGCTCCAGGCT 484  
Db 329 TGAACCTCGTGGAAAGACCGACAGCCGATCGTGGTTCGCTTCCATCGCCCGCA 388  
Qy 485 CTTCCATGAGTGTGATGGCCCCATGATCTTATACCGCGCTGAATGTAGCGATCAACA 544  
Db 389 CCGCATGTCCGCGACGGCATGTCAACCTGTACAAACGCGCTGGCGTGGCCGACCA 448  
Qy 545 AAGCTCTTACTAACAAAGGAGTGTGATTTGTATGAACGATGATTCACGCCGCCAGAG 604  
Db 449 AGGACTCGCGCGCAAGGGCGTGTGTGACCATGAACGACGAGATCCAGTCCGGGGTG 508  
Qy 605 AAGGACCAAGCTCAACACCAACCGAGTCAATGATTTGTTCCGCCCAACACAGTTAAA 664  
Db 509 ACGTGAGCAAGTCGATCAACATCAAGACCGAAGCTTC---AAGAGCGCTTGGGGCCGC 565  
Qy 665 TCGGCACAGTCTATTATGGCAAGTCGAGTATTTTCACTCAATCGTTTCGACCTCACACCC 724  
Db 566 TGGGATGCTGTGGAGGCAAGTGTACTGTTCCGCTCGCGCCGCGCAAGCGCACACGG 625  
Qy 725 TTGCAAGTGTGATTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTCGATATTTTACG 784  
Db 626 TCACTCCGAGTTCGACATCAAGCAGATCAGCAGCTGCCCGAGGTGACATCGCTACA 685  
Qy 785 CTCACCCCGATGATCTGATGTTTGTAGTCAATGAGCCCTTCAGCAGGAGCGCAAGGAA 844  
Db 686 GCTATGGCAACGTCACACACCGGCTTACAGGCGCTGGCACAGAACGCGCAAGCGC 745  
Qy 845 TCATCCATGCGAGCATGGCAATGGAAACCCCTTTCCCTTTGACTCAAAATGCTTTGAAA 904  
Db 746 TGATCCATCGCGCACCGGCAATGGCTCGGTCTCGTGGGGTGGTGGCCGCTGCAGG 805  
Qy 905 AAGCAGCCAAATCAGCGGTAGTGTGCTGCAAGCTCTAGAGTGGGCGTGGTCCACCA 964  
Db 806 AGCTGCGCAAGAACGCGCTGCAGATCATTCGTTCTGTCACGTCACAGGCGGTTTCGTGC 865  
Qy 965 CCCAAGAGGCTGAGTGGATGATAGAAACTTGTGTTTGTGCTACAGAGAGTCTCAACC 1024  
Db 866 TCGTTAACCGCGAGCAGCCCGCAGCAAGAACGACTGGTTCGTCGCGCCACGACCTGAAC 925  
Qy 1025 CTCAAAAAGCCAGAGTGTCTTTATGTTAGCCCTTCAACAACTAGTGTAGAGAGCGCA 1084  
Db 926 CGCAGAGGCCCGCATCTCTGGCGATGTTGGCAATGACCAAGACCCAGGACAGCAAGAGC 985  
Qy 1085 TCAAAAAGATCTTCTCACCCTATTAAT 1111  
Db 986 TGCAGCGCATTTTCTGGGATACTGAT 1012

RESULT 6

US-08-050-482A-1  
; Sequence 1, Application US/08050482A  
; Patent No. 6312939  
; GENERAL INFORMATION:  
; APPLICANT: ROBERTS, Joseph  
; APPLICANT: MACALLISTER, Thomas W.  
; APPLICANT: SETHURAMAN, Natarajan  
; APPLICANT: FREEMAN, Abbie G.  
; TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND  
; TITLE OF INVENTION: ITS USE IN ANTIVIRAL AND ANTICANCER THERAPY  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY & LARDNER  
; STREET: 3000 K Street, N.W.  
; CITY: Washington  
; STATE: D.C.

COUNTRY: U.S.A.  
 ZIP: 20007-5109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/050,482A  
 FILING DATE: 25-Apr-1995  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO PCT/US92/10421  
 FILING DATE: 04-DEC-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bent, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 023032/0106  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 672-5300  
 TELEFAX: (202) 672-5399  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1017 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..1011  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-08-050-482A-1

Query Match 18.6%; Score 210.8; DB 4; Length 1017;  
 Best Local Similarity 52.9%; Pred. No. 8.8e-60;  
 Matches 524; Conservative 0; Mismatches 457; Indels 9; Gaps 3;  
 QY 128 CCCAAGTACATCTAGCCACAGGAGGACCATCGCTGGTTCGGGGAATCTAGCGTCA 187  
 Db 29 CCAAGTGTGATCTCGCCACCGCGGCACATCGCGCGCTGGCGCAGCGCGGCA 88  
 QY 188 AGAG---TAGCTACTGTCTGGAGCAGTCACTGTTGATTAAGCTTCTTGCAGCGCTCCCTG 244  
 Db 89 ACAGCGCACCTACAGGCTGCCAAGGTGCGCTCGACAGCTGATTCGCGCGCTGCGG 149  
 QY 245 CCATCAACGACCTAGCCACCATCAAGGTTGAAGATCTCAAGCATTTGGTCCCAAGAGA 304  
 Db 149 AGCTGGCGGACCTGGCCCAATGTGCGCGCGAGCAGGTGATGCAGATCGCTCCGAAAGCA 208  
 QY 305 TGACGGGTAAAGTGTGGCTTAAACTAGCCAAAGCGTGTCAATGAGCTCTCGCCCAAAAG 364  
 Db 209 TCACCAACGACGACCTCTCAAGCTGGGCAAGCGGTGGCGGAGCTGGCGCAGCAATG 268  
 QY 365 AGACCGAAGCGGTGATCATACCCATGGAATCTATACCGCGTGAATGTAGCGATCAACA 424  
 Db 269 ACCTGATGGCATCGTCTATCACCATGGCACCCGACACCTTGGAGAAACCGCTACTTTT 328  
 QY 425 TCAACTTCAGCGTGAAGAGCAAAACCTGTCTGCTTGTAGCGGCGATCGTCCAGGCT 484  
 Db 329 TGAACCTCTGGAAAGACGACAGCCGATCGTCTGTCGCTGCTCCATGCGCGCGGCA 388  
 QY 485 CTTCCATGATGTGATGCGCCCATGAATCTTATACCGCGTGAATGTAGCGATCAACA 544  
 Db 389 CGGCCATGTCCGCGGACGGCATGCTCAACTGTGTACACCGCGTGGCGGTGGCGAGCA 448  
 QY 545 AAGCCTCTACTAACAAGAGGTGGTGTATGTATGACATGATGATCAAGCGCGCAGAG 604  
 Db 449 AGGACTCGCGGCAAGGCGGTGCTGGTGAACCATGAACGACGAGATCCAGTCCGCGGTG 508  
 QY 605 AAGGACCAAGTCAACACCGCAGTCAATGCAATTTGCTTCGCCCAACACAGGTAAA 664  
 Db 509 ACGTGAGCAAGTCGATCAACATCAAGACCGAAGCCTTC---AAGAGCGCTGGGCGCGC 565

QY 665 TCGGCACAGTCTATTATGCGAAAGTCGAGTATTTCATCAATCGTTTCGACCTCACACCC 724  
 Db 566 TGGCATGTGTGGTGAAGCAAGTCTACTGTGTTCCGCTTGCCTGCGGCAAGCGCACACGG 625  
 QY 725 TTGCAAGTCAAGTTGATATAGCAAAATCGAAGAACTCCCCAGAGTCGATATTCTTTACG 784  
 Db 626 TCACTCCGAGTTCGACATCAAGCAGATCAGCAGCCTGCCAGAGTGGACATCGCCTACA 685  
 QY 785 CTCACCCCGATGATCTGATGTTTATAGTCAATGACAGCCCTTCAGGCGAGGACCAAGGAA 844  
 Db 686 GCTATGGCAACGTCAACGACAGCGCTCAAGCGCCCTGGCACAGAAGCGCGCAAGCGC 745  
 QY 845 TCATCATGACGAGTGGCAATGGGAACCTTTTCCCTTTTGAATCAAAATGCTCTTGA 904  
 Db 746 TGATCCATGCCGCGCACCGGCAATGGTCTGCTGCGGGTGGTGGCAGCCCTCAGG 805  
 QY 905 AAGCAGCCAAATCAGGCGTAGTCTGCTGCGTTCGAGAGCTCTAGAGTGGGC---AGTGGTTCA 961  
 Db 806 AGCTCGGCAAGAACCGCGTGCAGATCATTTGCTTCCACCTCAACAGGCGGTTTCG 865  
 QY 962 CCACCAAGAGGCTGAAGTGGATGATAAGAAACTTTGTTTGTGCTACAGAGATCTCA 1021  
 Db 866 TGCTGCTGAACCGCGAGCAGCCCGACGACAGCAAGACGATGGTCTGGCCACGACCTGA 925  
 QY 1022 ACCCTCAAAAAGCCAGAGTCTTTTATGTTAGCCCTCACCAAACTAGTATAGAGAGG 1081  
 Db 926 ACCCGCAGAAGCGCCGATCCTGGCGATGTTGGCAATGACCAAGACCCAGGACCAAGG 985  
 QY 1082 CGATCCAAAAGATCTTCTCCACCTATTAAAT 1111  
 Db 986 AGCTGCAGCGCATTTTCTGGGAATCTGAT 1015

## RESULT 7

US-09-252-991A-5793  
 ; Sequence 5793, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 5793  
 ; LENGTH: 1182  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-5793

Query Match 17.7%; Score 200.4; DB 4; Length 1182;  
 Best Local Similarity 52.5%; Pred. No. 2.8e-56;  
 Matches 513; Conservative 0; Mismatches 456; Indels 9; Gaps 3;  
 QY 128 CCCAAGTACATCTAGCCACAGGAGGACCATCGCTGGTTCGGGGAATCTAGCGTCA 187  
 Db 197 CCAAGTGTGATCTCTGGCCACCGCGGCACCATCGCGCGCGCGCGCGCGCGCGCA 256  
 QY 188 AGAG---TAGCTACTCTGCTGGAGCAGTCACCGTTGATTAAGCTTCTTGCAGCGTCCCTG 244  
 Db 257 ACAGCGCCACCTACACCGCGCCCAAGGTGCGCGTGCACAGTGTCTGGCCAGCGTCCCG 316  
 QY 245 CCATCAACGACCTAGCCACCATCAAGGTTGAAGATCTCAAGCATTTGGCTCCCAAGAGA 304  
 Db 317 AGCTCAAGGACATCGCCAAACGTCGCGCGCAACAGGTGTTCCAGATCGCTCCGGAAGCT 376  
 QY 305 TGACGGGTAAAGTGTGGCTTAACTAGCCAAAGCGTGTCAATGAGCTCTCTCGCCCAAAAG 364

Db	377	TCACCAACGAGAACCTGCTGGAACTGGCGAAGACCGTCCGAAAGCTGGCGGATAGCGAGC	436
Qy	365	AGACGAGACCGGTGATCATCACCCATGGAACTGACACCATGGAAGAGACCGCTTTCTTCC	424
Db	437	ACGTGACGGCATCGGTGATCACCACCGTACCGACACCTTGAAGAGACCGCTACTTCC	496
Qy	425	TCACCTTCACGGTGAAGAGCAAAAACCTGTCTGCGCTTGTAGCGCCATCGTCCAGGCT	484
Db	497	TGACCTTGGTCGAGCACACCGAGAAGCTATCTGTGTGGTTCGCTCGATCGCCCGGCA	556
Qy	485	CTTCCATGAGTGTGATGGCCCCATGAACTCTATTAACGGCGTGAATGTAGCATCAACA	544
Db	557	CCGCGATGTCGCGCACCGCATGCTCAACTGTACAAACGCGTGGCGGTGGCGCGACA	616
Qy	545	AAGCCTCTACTAAAGAGAGTGGTGAATGTGTGATGAACGATGAGATTCACGGCGCCAGAG	604
Db	617	AGTGGCACCGCGCAGGGCGTGTGATCAACCATGAACGACGAGATCTCTTCGGCGCGG	676
Qy	605	AAGCGACCAAGCTCAACACACCGCAGTCAATGCATTTGCTTCGCCCAACACAGGTAAAA	664
Db	677	ACGGACCAAGTGGTCAAATCAAGACCGAAGGTTCAAGAGCCCGT---GGGGCCCGC	733
Qy	665	TCGSCACAGTCTATTATGGCAAGTGGAGTATTTCATCAATCCGTTTCGACCTCAACCC	724
Db	734	TGGCGATGTTGGTCTCGAGGCAAGAGCTACTGGTTCCGCGCACCGGTGAAGCGGCACACGG	793
Qy	725	TTGCAAGTGAGTTTGATATTAGCAAAATCGAAGAACTCCCCAGAGTCGATATCTTTACG	784
Db	794	TGAATCCGAGTTCCGACATCAAGCAGATCTCCGCGCTGGCTCCGGTGGAAATCGCTTACA	853
Qy	785	CTCACCCCGATGATCTATGATGTTTTAGTCAATGCAGCCCTTCAGGCAGGAGCCAAAGGAA	844
Db	854	GCTACGGCAAGCTCAGCGCACACCGCCTAAGAGCCCTGGCCCGCGCGTGGTCCGACCTGCAGG	913
Qy	845	TCATCTCATGAGGATGGCAATGGGAACCCCTTTCCCTTTTGACTCAAAATGCTCTTGA	904
Db	914	TCATCATTCGCGCACCGCAACCGCTCGGTGGCCCGCGCGTGGTCCGACCTGCAGG	973
Qy	905	AAGCAGCGCAATACAGGCGTAGTCTCGCTCGAAGCTCTTAGAGT---GGGCAGTGGTTCCA	961
Db	974	AACTGCGCAAGCAGGGCGTGACAGATCATTCGCTCCTCGCACGTCAATGCCCGCGGCTCG	1033
Qy	962	CCACCCAGAGCGTGAAGTGAATGAAGAACTTGGTTTTGTGGCTACAGAGAGTCTCA	1021
Db	1034	TCCTGCGCAACCGCGAGCAGCCGACGACGAAGAAGACGACTGGATCGTCGCCACGACCTGA	1093
Qy	1022	ACCTCTCAAAAACGACAGTGTCTTATTGTTAGCGCTTCAACAAATCTAGTGATAGAGAGG	1081
Db	1094	ACCCGCAAGCGCGCATCTTGGCCGCGTTCGCGATGACCAAGACCCAGGACAGCAAGG	1153
Qy	1082	CGATCAAAAAGATCTTCT	1099
Db	1154	AGCTGCAACGGATCTTCT	1171

## RESULT 8

RESOLUTION 8  
US-09-252-991A-5877/c  
; Sequence 5877, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 5877

	Query Match	17.7%	Score 200.4	DB 4	Length 1335
	Best Local Similarity	52.5%	Pred. No. 3e-56		
	Matches 513	Conservative 0	Mismatches 456	Indels 9	Gaps 3
Qy	128	CCCAAGTGA	CTATCTCCTAGGCACAGAGAGGACCATCGTCTGGTTCTGGGGGAATCTAGAGTCA	187	
Db	1211	CCAACTGGT	TCATCTCTGGCCACCGCGGACCATCGTCCGCGCGCGCGCGCAGCGCGCCCA	1152	
Qy	188	AGAG---	TAGTACTCTGCTGGAGCAGTCAACGTTTGATTAAGCTTTTGACGCGCTCCCTG	244	
Db	1151	ACAGCGCA	CCCTACACCGCGCCCAAGGTGCGGTTCGACAGTTGCTGGCAGGTCGCCG	1092	
Qy	245	CCATCAACG	ACCTAGGCACCATCAAGGGTGAAACAGATCTCAAGCATTGGCTCCCAAGAGA	304	
Db	1091	AGTCTAAG	GACATCGCCACGTCGCGGGGAAACAGGTGTTCCAGATCGCTCGGAAGCT	1032	
Qy	305	TGACGGGT	AAGTGTGGCTTAAACTAGCCAGCGTGTCAATGAGCTCTCTGCGCCCAAAAG	364	
Db	1031	TCACCAAC	GAGAACCTGCTGGAATCGGGCAAGCCGTCGCAAGCTTGGCCGATAGCGAGC	972	
Qy	365	AGACCGA	AGCGGTGATCATCACCATTGGAATCGACACCATGGAAGAGACCGCTTTCTTCC	424	
Db	971	ACGTGCA	GGCATCGTGATCACCACCGTACCACACCTGGAAGAGACCGCCTACTTCC	912	
Qy	425	TCAACTC	ACGGTGAAAAGCAAAACCTGTCTGCCTTTGAGCGGCATCGCTCCAGGCT	484	
Db	911	TGACCTG	TCTGAGCACACCGAGAGCCTATCTGTGTGTGCTCGCTCGATCGCCCGGCA	852	
Qy	485	CTTCCAT	GATGTGATGGCCCATGAATCTCTATAAGCCCGTGAATGTAGCGATCAACA	544	
Db	851	CCGCGAT	GTCCGCGACGGCATGCTCAACCTGTACACGCCGTGGCGGTGGCGCGGACA	792	
Qy	545	AAGCCTC	TACTAACAAAGGAGTGGTGATTGTGATGAACGATGAGATTCACGCGCCAGAG	604	
Db	791	AGTCGGC	ACCGGCAAGGGCGTCTGATCACCATGAACAGAGATCTCTCCGGCCGCG	732	
Qy	605	AAGCGAC	CAAGCTCAACACCAACCGCAGTCAATGCTTTGCTTCGCCCAACACAGGTA	664	
Db	731	ACGCGAC	AGATGGTCAACATCAAGACCGAAGCGTTCAAGAGCCCGT---GGGCGCCGC	675	
Qy	665	TCGGCAG	ATCTATTATGGCMAAGTCAGATATTCTACTCAATCCGTTGACCTCACACC	724	
Db	674	TGGGCAT	GTTGGTTCAGGCGCAAGAGCTACTGGTTTCGCGCACCGGTGAAGCGGCACACG	615	
Qy	725	TTGCAAG	TGAGTTTGATATTAGCAAAATCGAAGAACTCCCGAGAGTCGATATCTTTACG	784	
Db	614	TGAATCG	AGTTTCGATCAATCAAGCAGATCTCCGCGCTGGCTCCGTTGGAATCGCCTACA	555	
Qy	785	CTCACCC	CGATGACTGATGTTTATGTCATATGACGCCCTTCAGCGCAGGAGCCAAAGAA	844	
Db	554	GCTACGG	CAACGTCAGGACACCGCTACAAGGCCCTGGCCAGCGCGCGCCCAAGGCGA	495	
Qy	845	TCATCCAT	GAGGATGGGCAATGGGAACCTTTCCCTTTGACTCAAAATGCTCTTGAA	904	
Db	494	TCATCCAT	GCGGACCGGCAACGGCTCGGTGCCCGCGCGGTGGTGCAGCCCTTGACG	435	
Qy	905	AAGCAGC	CAAAATCAGGCGTAGTCTGTCGCTCGAAGCTCTAGAGT---GGGCAAGTGTTCCA	961	
Db	434	AATCGCG	AAGAGGGCGTGCAGATCATCCGCTCTCGCAGTCAATGCGCGCGGCTTCG	375	
Qy	962	CCACCCAA	GAGCTGAAGTGGATGAAGAACTTTGTTGTGGCTACAGAGAGTCTCA	1021	
Db	374	TCCTCGC	CAACCGCAGCAGCGACGCAAGAACGACTGGATCGTCGCCCAACGACCTGA	315	
Qy	1022	ACCTCAAA	AGCCAGAGTGCTTTTATGTTAGCCCTCACCAAACTAGTGTATGAGAGG	1081	
Db	314	ACCCGCA	GAAGCGCGCATCTCTGGCGGGGTTCGCGATGACCAAGACCCAGGACGACGAGG	255	



Query Match	8.9%;	Score 101.4;	DB 4;	Length 552;
Best Local Similarity	84.4%;	Pred. No. 1.5e-23;		
Matches 114;	Conservative 0;	Mismatches 21;	Indels 0;	Gaps 0;
QY	1	ATGGGAGAGCCATCATCATCATCATCATATGAGAGGGGCTGTGTGGCGCGGAGCCAT	60	
DB	1	ATGGGAGAGCCATCATCATCATCATCATGAGAGGGGCTGTGTGGCGCGGAGCCAT	60	
QY	61	ATGGCTAGCATGACTGTGTGGACGCAAAATGGGTGGGGATCCAGCGGCTCTGTTTTCATG	120	
DB	61	ATGGCTAGCATGACTGTGTGGACGCAAAATGGGTGGGGATCGCTGGGCTTTGTT	120	
QY	121	GCTAAACCCCAAGTG	135	
DB	121	GGTAACAAAGCAGTG	135	

RESULT 12  
 US-09-252-991A-5837/C  
 ; Sequence 5837, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 5837  
 ; LENGTH: 270  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-5837  
  
 Query Match 7.1%; Score 80; DB 4; Length 270;  
 Best Local Similarity 59.8%; Pred. No. 1.3e-16;  
 Matches 134; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
  
 QY 418 TTCCTCTCAACCTCACGGTGAAGAGCCAAAACCTGTCTGCCTTGTAGGCGCCATGCGT 477  
 Db 269 TACTTCTGACCTTGTCTGAGCACACCGAGAGCCCTATCGTGGTGGTCGGCTCGATGCGC 210  
  
 QY 478 CAGGCTCTTCATGAGTGCTGATGGCCCCCATGAATCTCTATTAACGCCGTGAATGAGCG 537  
 Db 209 CCGGGCACCGCATGTGCCCGCAGCGCATGCTCAACCTGTACAAACCCGGTGGCGGTGGCC 150  
  
 QY 538 ATCAACAAGCCCTCTACTAACAAAGGAGTGGTGATGTGATGAACGATGAGATTCACGCC 597  
 Db 149 GGGCAAGAATGCGCAGCGGCAAGGGCGTGCTGATCACCATGAACGACGAGATCCTCTCC 90  
  
 QY 598 GCACAGAGACGACCAAGCTCAACACACCGCAGTCAATGCATT 641  
 Db 89 GGGCGGCGAGCGAGCAAGATGGTCAACATCAAGACCGAGCGTT 46  
  
 RESULT 13  
 US-09-107-532A-1293  
 ; Sequence 1293, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD-ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107,532A  
 ; FILING DATE: 30-Jun-1998  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085,598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:

```

; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1293:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1005 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8) LOCATION 1...1005
; SEQUENCE DESCRIPTION: SEQ ID NO: 1293:
US-09-107-532A-1293

```

```

Query Match      6.8%; Score 76.6; DB 4; Length 1005;
Best Local Similarity 53.5%; Pred. No. 3.9e-15;
Matches 160; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 376 GTGATCATCACCCATGGAACTGACACCATGGAAGAGACCGCTTTCTTCTCAACTCAG 435
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 277 GTGGTGATCACACCGGTACAGTACATTAGAGAACTGCTTTTCTTGTATACACG 336
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 GTGAAGCCAAAACCTGCTGCTGTAGCGCCATGCGTCCAGGCTCTTCCATGAGT 495
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 337 ATTGGCAACAATTACCAATTTGTTTGAATGGGCAATGCTTCAAGTAATGAATTGA 396
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 GCTGATGCCCCCATGAATCTCTATAACGCGGTGATGTAGCGATCAACAAGCCCTTACT 555
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 397 AGTACGCGCTTTATATTTTGAAGTGCAGTTCGTGTGCGAAGTTGTGAAGAGCGCTA 456
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 AACAAAGAGTGTGATGTGATGAACGATGAGATTCACGCGCCGAGAGAGCCACCAAG 615
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 457 GATAAGGGGTTTTGGTGGTCATGAACGATGAGATCCATTCTGACGCTACGTGACAAAG 516
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 616 CTCACACACCGGAGTCAATGCTTTCGCGCCACACAGTAAATCGGCACAGT 674
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 ACCATACCAAAATGTTGCTACTTTCCGAACACCTACATTGGGTGCCATCGGCTTAGT 575
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 14
US-08-961-527-265
; Sequence 265, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 265:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2231 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-265

Query Match      6.5%; Score 74; DB 4; Length 2231;
Best Local Similarity 51.9%; Pred. No. 4.7e-14;
Matches 190; Conservative 0; Mismatches 175; Indels 1; Gaps 1;

QY 369 CGAAGCCGTGATCATCACTCCATGGAACCTGACACCATGGAAGAGACCGCTTTCTTCTCAA 428
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 429 CCTCAGCGTGAAAGCCAAACCTGTCTGCTTGTAGGCGCCATGCGTCCAGGCTCTTC 488
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1137 TACCATGGAAGTTCCCATATATGCTTCTTAACAGGAGCCATGGTACTC-CAATGA 1195
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 489 CATGAGTGTGATGGCCCCCATGATCTCTATAACGCGGTGATGTAGGATCAACAAGC 548
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1196 GCTCGGTAGTGTGTTTATATTAATTAACGTGCTTACGAGTGCCAGCGATGACAG 1255
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 549 CTCTACTAAACAAAGGAGTGTGATGTGATGAACGATGAGATTCACGCGCCGAGAGAGC 608
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1256 GGCTGCTGACAAAGGAGTTTGGTGTGTTATGACGATGAATCCACGTCGCAAGTATGT 1315
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 609 GACCAGCTCAACACACACCGCAGTCAATGCTTTCGCGCCACACAGAGTAAATCGG 668
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1316 CACCAAAACACATACGACTAATGTGACAGCTTCCAGACTCCAAACACATGCCCCCTGG 1375
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 669 CACAGTCTATTTATGGCAAGTCGAGTATTTCACTCAATCGTTTCGACTCACACCTTGC 728
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1376 TCTCATCAUGAAACAGGAATCTCTACTTCAAAACAGCTGAACCTCGTGTGCTTTGA 1435
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 729 AAGTGA 734
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1436 CCTTGA 1441
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-09-613-303-11
; Sequence 11, Application US/09613303
; Patent No. 6495347
; GENERAL INFORMATION:
; APPLICANT: Siegel, Marvin
; APPLICANT: Chu, N. Randall
; APPLICANT: Mizzen, Lee A.
; TITLE OF INVENTION: INDUCTION OF A TH1-LIKE RESPONSE IN VITRO
; FILE REFERENCE: 12071/002001
; CURRENT APPLICATION NUMBER: US/09/613,303
; PRIOR FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: US 60/143,757
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion sequence
; NAME/KEY: CDS

```





Thu Aug 28 09:00:59 2003

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 02:34:35 ; Search time 843.661 Seconds  
(without alignments)  
3081.078 Million cell updates/sec

Title: US-09-937-982-3  
 Perfect score: 1133  
 Sequence: 1 atgggaagcagcatcatca.....aagaagggaatctcttcac 1133

Scoring table: IDENTITY\_NUC  
Gapop 10.0 ; Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues  
Total number of hits satisfying chosen parameters: 3067400

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 su

Database : Published Applications NA: \*

11: /cgm2/6/ptodata/2/pubnpa/US07\_PUBCOMB.seq.\*  
12: /cgm2/6/ptodata/2/pubnpa/PCT\_NEW\_PUB.seq.\*  
13: /cgm2/6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq.\*  
14: /cgm2/6/ptodata/2/pubnpa/US06\_PUBCOMB.seq.\*  
15: /cgm2/6/ptodata/2/pubnpa/US07\_NEW\_PUB.seq.\*  
16: /cgm2/6/ptodata/2/pubnpa/PCTUS\_PUBCOMB.seq.\*  
17: /cgm2/6/ptodata/2/pubnpa/US08\_NEW\_PUB.seq.\*  
18: /cgm2/6/ptodata/2/pubnpa/US08\_PUBCOMB.seq.\*  
19: /cgm2/6/ptodata/2/pubnpa/US09A\_PUBCOMB.seq.\*  
20: /cgm2/6/ptodata/2/pubnpa/US09B\_PUBCOMB.seq.\*  
21: /cgm2/6/ptodata/2/pubnpa/US09C\_PUBCOMB.seq.\*  
22: /cgm2/6/ptodata/2/pubnpa/US09\_NEW\_PUB.seq.\*  
23: /cgm2/6/ptodata/2/pubnpa/US10A\_PUBCOMB.seq.\*  
24: /cgm2/6/ptodata/2/pubnpa/US10B\_PUBCOMB.seq.\*  
25: /cgm2/6/ptodata/2/pubnpa/US10\_NEW\_PUB.seq.\*  
26: /cgm2/6/ptodata/2/pubnpa/US10\_PUBCOMB.seq.\*  
27: /cgm2/6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq.\*  
28: /cgm2/6/ptodata/2/pubnpa/US06\_PUBCOMB.seq.\*  
29: /cgm2/6/ptodata/2/pubnpa/US06\_NEW\_PUB.seq.\*  
30: /cgm2/6/ptodata/2/pubnpa/US06\_PUBCOMB.seq.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	129.8	99.7	1133	10	US-09-773-260-3	Sequence 3, Appli
2	296.4	26.2	1830121	14	US-10-328-980-1	Sequence 1, Appli
3	259.2	22.9	1196	12	US-09-882-227-83	Sequence 83, Appli
4	215.8	19.0	1014	9	US-09-842-628-1	Sequence 1, Appli
5	103.6	9.1	543	10	US-09-876-766A-34	Sequence 34, Appli
6	103.6	9.1	543	10	US-09-876-766A-34	Sequence 34, Appli
7	102	9.0	543	10	US-09-876-348A-18	Sequence 18, Appli
8	102	9.0	543	10	US-09-876-348A-22	Sequence 22, Appli
9	102	9.0	543	10	US-09-876-348A-30	Sequence 30, Appli
10	102	9.0	543	10	US-09-876-348A-38	Sequence 38, Appli
11	102	9.0	543	10	US-09-876-796A-18	Sequence 18, Appli
12	102	9.0	543	10	US-09-876-796A-22	Sequence 22, Appli
13	102	9.0	543	10	US-09-876-796A-30	Sequence 30, Appli
14	102	9.0	543	10	US-09-876-796A-38	Sequence 38, Appli
15	101.6	9.0	1436	12	US-10-228-063-43	Sequence 43, Appli
16	100.4	8.9	681	10	US-09-876-348A-16	Sequence 16, Appli

17	100.4	8.9	681	10	US-09-876-348A-28	Sequence 28, Appl
18	100.4	8.9	681	10	US-09-876-348A-36	Sequence 36, Appl
19	100.4	8.9	681	10	US-09-876-796A-16	Sequence 16, Appl
20	100.4	8.9	681	10	US-09-876-796A-28	Sequence 28, Appl
21	100.4	8.9	681	10	US-09-876-796A-36	Sequence 36, Appl
22	100.4	8.9	681	10	US-09-876-348A-20	Sequence 20, Appl
23	100.4	8.9	682	10	US-09-876-348A-32	Sequence 32, Appl
24	100.4	8.9	682	10	US-09-876-796A-20	Sequence 20, Appl
25	100.4	8.9	682	10	US-09-876-796A-32	Sequence 32, Appl
26	100.4	8.9	682	10	US-09-876-348A-24	Sequence 24, Appl
27	100.4	8.9	776	10	US-09-876-796A-24	Sequence 24, Appl
28	98.4	8.7	1435	12	US-10-228-063-41	Sequence 41, Appl
29	96.4	8.5	543	10	US-09-876-348A-26	Sequence 26, Appl
30	96.4	8.5	543	10	US-09-876-796A-26	Sequence 26, Appl
31	94.2	8.3	900	10	US-09-850-677A-1	Sequence 1, Appl
32	93.4	8.2	5312	12	US-10-263-103-35	Sequence 35, Appl
33	72.6	6.4	366	14	US-10-267-311-11	Sequence 11, Appl
34	70.4	6.2	1230	14	US-10-267-311-54	Sequence 54, Appl
35	68.4	6.0	6682	10	US-09-070-927A-56	Sequence 56, Appl
36	66.2	5.8	2847	14	US-10-267-311-20	Sequence 20, Appl
37	65.6	5.8	6061	12	US-10-314-861-10	Sequence 10, Appl
38	65.6	5.8	6100	12	US-10-314-861-36	Sequence 36, Appl
39	65.6	5.8	6115	12	US-10-314-861-38	Sequence 34, Appl
40	65.6	5.8	6121	12	US-10-314-861-38	Sequence 38, Appl
41	65.6	5.8	6133	12	US-10-314-861-32	Sequence 32, Appl
42	65.6	5.8	6151	12	US-10-314-861-30	Sequence 30, Appl
43	65.6	5.8	6169	12	US-10-314-861-28	Sequence 28, Appl
44	65.6	5.8	6248	12	US-10-314-861-15	Sequence 15, Appl
45	65.6	5.8	1289	14	US-10-269-557-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
 US-09-773-260-3  
 ; Sequence 3, Application US/09773260  
 ; Patent No. US20020102251A1  
 GENERAL INFORMATION:  
 APPLICANT: Donald L Durden  
 TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES  
 ; ASPARAGINASE IN THE TREATMENT OF HUMAN  
 ; HEMATOLOGIC AND AUTOIMMUNE DISEASE  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071-2066  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Fastseq for Windows 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/773,260  
 FILING DATE: 31-Jan-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/094,435  
 FILING DATE: 1998-06-09  
 APPLICATION NUMBER: <Unknown>  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Warburg, Richard J.  
 REGISTRATION NUMBER: 32,327  
 REFERENCE/DOCKET NUMBER: 234/274  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600

Query Match	99.7%;	Score 1129.8;	DB 10;	Length 1133;
Best local Similarity	99.8%;	Pred. No. 0;		
Matches 1131;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1	ATGGGCGAGCAGCCATCATCATCATCATAGCAGCGCCTGGTGC	CGCGCGGAGCCAT	60
Db	1	ATGGGCGAGCAGCCATCATCATCATCATAGCAGCGCCTGGTGC	CGCGCGGAGCCAT	60
QY	61	ATGGCTAGCATGACTGGTGGACAGCAAAATGGGTGCGGATCCAGCGCCTCTGTTTGTATG	120	
Db	61	ATGGCTAGCATGACTGGTGGACAGCAAAATGGGTGCGGATCCAGCGCCTCTGTTTGTATG	120	
QY	121	GCTAAACCCCAAGTACTCTCTGCTGGAGCAGTCAACCGTTGATAAAGCTTCTTCGAGCGCTC	240	
Db	121	GCTAAACCCCAAGTACTCTCTGCTGGAGCAGTCAACCGTTGATAAAGCTTCTTCGAGCGCTC	240	
QY	181	AGCGTCAAGAGTAGTACTCTCTGCTGGAGCAGTCAACCGTTGATAAAGCTTCTTCGAGCGCTC	240	
Db	181	AGCGTCAAGAGTAGTACTCTCTGCTGGAGCAGTCAACCGTTGATAAAGCTTCTTCGAGCGCTC	240	
QY	241	CTTGCCATCAACGACCTAGCCACATCAAGGGTGAAACAGATCTCAAGCATTTGGTCCCAA	300	
Db	241	CTTGCCATCAACGACCTAGCCACATCAAGGGTGAAACAGATCTCAAGCATTTGGTCCCAA	300	
QY	301	GAGATCAGGGTAAAGTGTGGCTTAACTAGCCACAGCGTGTCAATGAGTCTCTCGCCAA	360	
Db	301	GAGATCAGGGTAAAGTGTGGCTTAACTAGCCACAGCGTGTCAATGAGTCTCTCGCCAA	360	
QY	361	AAAGAGACCGAAGCGGTGATCATACCCATGGAACTGACACCATGGAAGAGACCGCTTTC	420	
Db	361	AAAGAGACCGAAGCGGTGATCATACCCATGGAACTGACACCATGGAAGAGACCGCTTTC	420	
QY	421	TTCTTCAACTCAGCGTGAAGAGCCAAACCTGTCTGCTGTGTAGGCGCCATGGGTCCA	480	
Db	421	TTCTTCAACTCAGCGTGAAGAGCCAAACCTGTCTGCTGTGTAGGCGCCATGGGTCCA	480	
QY	481	GGCTCTTCCATGAGTGTGATGGCCCGCCATGAATCTCTATTAACCGCGTCAATGTAGCGATC	540	
Db	481	GGCTCTTCCATGAGTGTGATGGCCCGCCATGAATCTCTATTAACCGCGTCAATGTAGCGATC	540	
QY	541	AACAAGCGCTCTACTAAACAAAGAGTGTGATTTGATGAACAGATGAGATCAACGCGCC	600	
Db	541	AACAAGCGCTCTACTAAACAAAGAGTGTGATTTGATGAACAGATGAGATCAACGCGCC	600	
QY	601	AGAGAGCGCACCAAGCTCAACACCAACCGCAGTCAATGATTTGCTTCCGCCAAACACAGGT	660	
Db	601	AGAGAGCGCACCAAGCTCAACACCAACCGCAGTCAATGATTTGCTTCCGCCAAACACAGGT	660	
QY	661	AAATTCGGCAGTCTATTATGGCAAGTCSAGTATTTCACTCAATCCGTTTCGACCTCAC	720	
Db	661	AAATTCGGCAGTCTATTATGGCAAGTCSAGTATTTCACTCAATCCGTTTCGACCTCAC	720	
QY	721	ACCTTTCGCAAGTGTGATTTAGCAGAAATCCAGNACTCCCGCAGTGCATATTCCT	780	
Db	721	ACCTTTCGCAAGTGTGATTTAGCAGAAATCCAGNACTCCCGCAGTGCATATTCCT	780	
QY	781	TACGCTCACCCGATGATCTGATGTTTGTAGTCAATGCAGCCCTTCAGGAGGAGCCAA	840	
Db	781	TACGCTCACCCGATGATCTGATGTTTGTAGTCAATGCAGCCCTTCAGGAGGAGCCAA	840	
QY	841	GGATCATCCATGAGCAGTGGGCAATGGGNAACCCCTTCCCTTGTGACTCAAAATGCTCTT	900	
Db	841	GGATCATCCATGAGCAGTGGGCAATGGGNAACCCCTTCCCTTGTGACTCAAAATGCTCTT	900	

```

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c

```



697	AGGGAGCGTGATATTGGCCAAACGGCGCTATTACATGCAGCGCTTTGAGAAAAACACACCAC	756
726	TGCAAGTGAAGTTGATATTAGCAAAATCGAA---GAACTCCCGAGAGTCGATATCTTTTA	782
757	AGAGAGCGAAATTTTCCCTTTTCACAACTCAAAACCCCGCTGCCCTAAAGTCGATATTATTTTA	816
783	CGCTCACCCCGATGATACTGATGTTTTTGTAGTCAATGCAGCGCCTTCAGGCGAGGACCCAAAGG	842
817	CACGCATGCTGGCATGACCCCTGATTTATTTCAGCGAGCCTTAACTCGCATGCAAAAGG	876
843	AATCATCCATGCAGGCGATGGGCAATGGGAACCCCTTTCCCTTTGACTCAAAATGCTCTTGA	902
877	CGTTGTGATAGCGGGTGGTAATGGGAATGTGAGCGTGGGTTTTTAAAGCGGATGCA	936
903	AAAGCAGCGCAATCAGGGGTAGTCTGCTCGAAGCTCTAGAGTGGCGAGTGGTTCCAC	962
937	AGAAGCGAGCCAAATGGGGTGGTTATTGTTGCTCTTAGCAGGGTAAATAGCGGTCAGAT	996
963	CACCCAAAGAGGCTGAAGTGGATGAAGAAACTTGGTTTTGTGGCTACAGAGAGTCTCAA	1022
997	TA-----CTTCAGGCGAGATTGATGACAGGCGCTTCATCACAAGCGCAATTTAA	1047
1023	CCCTCAAAAGCCAGAGTGCTTCTTATGTTAGCGCTCACCAAAACTGATGATAGAGGCG	1082
1048	CCCCCAAAAGCTAGGCTGCTTTTACAACCTCGCTTTAACTAAAAACAATAATAAGAAAA	1107
1083	GATCCAAAAGATCTTCTCCACCTATTAA	1110
1108	AATCCAAGAAATGTTTGAAGAGTATTGA	1135

## RESULT 4

```

US-09-842-628-1
; Sequence 1, Application US/09842628
; Patent No. US20020064862A1
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: MACALLISTER, THOMAS W.
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: FREEMAN, ABBIE G.
; TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND ITS USE IN
; TITLE OF INVENTION: ANTIVIRAL AND ANTICANCER THERAPY
; FILE REFERENCE: 023032/0108
; CURRENT APPLICATION NUMBER: US/09/842,628
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 08/050,482
; PRIOR FILING DATE: 1995-04-25
; PRIOR APPLICATION NUMBER: PCT/US92/10421
; PRIOR FILING DATE: 1992-12-04
; PRIOR APPLICATION NUMBER: DE P 4140003.8
; PRIOR FILING DATE: 1991-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Pseudomonas sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1008)
; US-09-842-628-1

```

89	DB	89	AAGCGCCACCTTACGAGCGTCCCAAGGTTGGCGTGCACAAGCTGATTCCCGCGGTGCCGG	148
245	QY	245	CCATCAACGAGCTTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTTCCCAAGAGA	304
149	DB	149	AGCTGGCCGACCTGGCCCAATGTGCGCGCAGCAGGTGATCAGATCGCTTCGGAAGACA	208
305	QY	305	TCACGGGTAAAGTGTGGCTTAAACTAGCCAGCGTGTCAATAGACTCTCTCCGCCAAAAAG	364
209	DB	209	TCACCAACGAGCACTGCTCAAGCTTGGCAAGCGGTGGCCAGCAGCACTG	268
365	QY	365	AGACCGAAGCCGTGATCATCAACCCATGGAACCTGACACCATGGAAGAGACCGCTTCTTCC	424
269	DB	269	AGCTCGATGGCATCGTCAATCAACCCATGCAACCGACACCTCGGAAGAAACCGCTACTTTT	328
425	QY	425	TCAACTCAGCGTGAAAGCCAAACCACTGTCTGCTTGTAGGCCCATCGTCCAGCT	484
329	DB	329	TGAACCTCGTGGAAAGACCGACAAGCCGATCGTCGTGGTCCGTTCATGCGCCCGCA	388
485	QY	485	CTTCCATGAGTCTGATGCGCCCATGAATCTCTATAAGCCCGTGAATGTAGCATCAACA	544
389	DB	389	CGGCCATGTCGCGCAGCGCATGCTCAACCTGTACAACGCGTGGCCGTGGCAGCAACA	448
545	QY	545	AAGCCTCTACTAAACAAGAGTGGTATTGTGATGAACGATCAGATTTCACGCCGCAGAG	604
449	DB	449	AGGACTTCGCGCGCAAGCGCGTGTGGTGAACCATGAACGAGATCCAGTCCGGCGGTG	508
605	QY	605	AAGGACCAAGCTCAACACCCGCGAGTCAATGCAATTTGCTTCGCGCCAAACACAGGTAAAA	664
509	DB	509	AGTGAGCAAGTCGATCAACATCAAGCCGAGCCCTC---AAGAGCGCTGGGCGCCGC	565
665	QY	665	TCGGCACAGTCTATTATGTCGAAGTCGAGTATTTCACTCAATCCGTTTGCACCTCAACCC	724
566	DB	566	TGGCATGTGTGTGAAGCAAGTGGTACTGTGTTCCGCTGCGCGCAAGCGCCACACGG	625
725	QY	725	TTGCAAGTCAGTTTGATATTAGCAAAATCGAAGAACTCCCCAGAGTCGATATTCTTTACG	784
626	DB	626	TCAACTCCGAGTTTCGACATCAAGCAGATCAGCAGCTGCCCAGGTGGACATCGCCTACA	685
785	QY	785	CTCACCCCGATGATCTGATGTTTTAGTCAATTCGAGCCCTTCAGGCAGGAGCCAAAGAA	844
686	DB	686	GCTATTGGCAAGTCAACCGACCGGCCCTACAAGGCCCTGGCACAGAACCGCGCAAGCGC	745
845	QY	845	TCATCCATCAGGCATGGCAATGGAAACCTTTCCCTTTGACTCAAAATGCTCTTTGAAA	904
746	DB	746	TGATCATCGCGCACCGGCAATGGCTCGGTGCTGCGGGTGTGTCAGCCCTGCAGG	805
905	QY	905	AAGCAGCCAAATCAGGCGTAGTCGTGCTCGACTCGAAGCTCTAGAGTGGCAGTGGTTCCACCA	964
806	DB	806	AGCTGCGCAAGAACGCGGTGCAGATCATTTGTTTCGTCACTCAACAGGGCGGTTTCGTGC	865
965	QY	965	CCCAAGAGGCTGAAGTGGATGATAGAAACATTGGTTTTTGTGCTACAGAGAGTCTCAAC	1024
866	DB	866	TGGGTAAACCGCAGAGCCCGACGACAAGAACGACTGGGTGTGGGCCCGACGACCTGACC	925
1025	QY	1025	CTCAAAAGCCAGAGTGCCTTTTATGTATTAGCCCTCACCAAACTAGTGTATGAGAGCGCA	1084
926	DB	926	CGCAGAAGCCCGCATCTCTGGCGATGTTGGCAATGACCAAGACCCAGGACAGCAAGGAGC	985
1085	QY	1085	TCAAAAGATCTTCTCCACCTATTAT	1111
986	DB	986	TGCAGCGCATTTTCTGGGAATACTGAT	1012

## RESULT 5

RESULT 5  
US-09-876-348A-34  
; Sequence 34, Application US/09876348A  
; Patent No. US20020172951A1  
; GENERAL INFORMATION:  
; APPLICANT: Horwath, K. L. and Myers, K. L.  
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio  
; TITLE OF INVENTION: Antifeedz Proteins and Method for Assaying Activity.  
; FILE REFERENCE: RB-125-RI  
; CURRENT APPLICATION NUMBER: US/09/876,348A



```

RESULT 9
US-09-876-348A-30
; Sequence 30, Application US/09876348A
; Patent No. US20020172951A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Myers, K. L.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
; FILE REFERENCE: RB-125-R1
; CURRENT APPLICATION NUMBER: US/09/876,348A
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 30
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: His-tagged, Signal minus, Clone 3.4
US-09-876-348A-30

Query Match          9.0%; Score 102; DB 10; Length 543;
Best Local Similarity 88.1%; Pred. No. 3.8e-24;
Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1   ATGGGCAGCAGCCATCATCATCATCATCATCATGACGCGCTGGTGC CGCGCGCAGCCAT 60
         |||
Db       64   ATGGGCAGCAGCCATCATCATCATCATCATGACGCGCTGGTGC CGCGCGCAGCCAT 123
         |||

QY      61   ATGGCTAGCATGACTGGTGGACAGCAAAATGGGTGCGCGATCCAGCGGCTCTGTTTGATG 120
         |||
Db     124   ATGGCTAGCATGACTGGTGGACAGCAAAATGGGTGCGCGATCCCTCACCGACGACAGATA 183
         |||

QY     121   GCTAAA 126
         ||||
Db    184   CAGAAA 189

```

```

; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Myers, K. L.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
; FILE REFERENCE: RB-125-R1
; CURRENT APPLICATION NUMBER: US/09/876,348A
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 38
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: His-tagged, Signal minus, Clone 7.5
US-09-876-348A-38

Query Match          9.0%;   Score 102;   DB 10;   Length 543;
Best Local Similarity 88.1%;   Pred. No. 3.8e-24;
Matches 111;   Conservative 0;   Mismatches 15;   Indels 0;   Gaps 0

QY      1  ATGGGCACAGCCATCATCATCATCATATGACGAGCGCTGGTCCGCGGGCAGCCAT 60
          |||
DB      64  ATGGGCACAGCCATCATCATCATATGACGAGCGCTGGTCCGCGGGCAGCCAT 123
          |||

QY      61  ATGGCTAGCATGACTGGTGGACAGCAATATGGCTCGCGGATCCAGGCGCTCTCTTTTGTATG 120
          |||
DB     124  ATGGCTAGCATGACTGGTGGACAGCAATATGGCTCGCGGATCCCTCACCGACAAACAGATA 183
          |||

```

Qy	61	ATGGCTAGCATGCTGGTGGACAGCAAAATGGGTGCGGATCCAGCGCTCTGTTTTGATG	120
Db	124	ATGGCTAGCATGCTGGTGGACAGCAAAATGGGTGCGGATCCCTCACCGACGAACAGATA	183
Qy	121	GCTAAA	126
Db	184	CAGAAA	189

RESULT 13

```

US-09-876-796A-30
; Sequence 30, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; TITLE OF INVENTION: Antifeedz Proteins and Method for Assaying Activity.
; FILE REFERENCE: RH-125-SEQ
; CURRENT APPLICATION NUMBER: US/09/876, 796A
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 30
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; -OTHER INFORMATION: His-tagged, Signal minus, Clone 3.4
US-09-876-796A-30

```

Query Match	9.0%; Score 102; DB 10; Length 543;
Best Local Similarity	88.1%; Pred. No. 3.8e-24;
Matches 111; Conservative	0; Mismatches 15; Indels 0; Gaps 0;
QY	1 ATGGGCAGCAGCCATCATCATCATCATCATGACGGCGCTGGTCCGCGCGCGCAGCCAT 60
Db	64 ATGGGCAGCAGCCATCATCATCATCATCATGACGGCGCTGGTCCGCGCGCGCAGCCAT 123
QY	61 ATGGCTAGCATGACTGGTGACAGCAAAATGGTTCGGGATCCAGGCCCTCTCTTTTGATG 120
Db	124 ATGGCTAGCATGACTGGTGACAGCAAAATGGTTCGGGATCCCTCACCGACAACAGATA 183

RESULT 14

```

US-09-876-796A-38
; Sequence 38, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
; FILE REFERENCE: RB-125-SEO
; CURRENT APPLICATION NUMBER: US/09/876,796A
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 38
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: His-tagged, Signal minus, Clone 7.5
US-09-876-796A-38

```

Query Match 9.0%; Score 102; DB 10; Length 543;  
Best Local Similarity 88.1%; Pred. No. 3.8e-24;

	Matches	111;	Conservative	0;	Mismatches	15;	Indels	0;	Gaps	0;
Qy	1	ATGGGCAGCAGCCATCATCATCATCATATGACAGCGGCTGTGTCGCGCGGCAGCCAT	60							
Db	64	ATGGGCAGCAGCCATCATCATCATCATATGACAGCGGCTGTGTCGCGCGGCAGCCAT	123							
Qy	61	ATGGCTAGCATGACTGGTGGACAGCAAAATGGGTGCGGATCCAGCGGCTCTGTTTTCATG	120							
Db	124	ATGGCTAGCATGACTGGTGGACAGCAAAATGGGTGCGGATCCCTACCGACGAACAGATA	183							
Qy	121	GCTAAA	126							
Db	184	CAGAAA	189							

## RESULT 15

```

US-10-228-063-43
; Sequence 43, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lananhan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1436
; TYPE: DNA
; ORGANISM: Thermotoga neapolitana
US-10-228-063-43

```

	Query Match	9.0%;	Score 101.6;	DB 12;	Length 1436;
	Best Local Similarity	76.2%;	Pred. No. 9.7e-24;		
	Matches 125;	Conservative 0;	Mismatches 39;	Indels 0;	Gaps 0;
QY	1	ATGGGCAGCAGCCATCATCATCATATAGCAGCGGCTTGGTCCGCGCGGACGCCAT	60		
Db	1	ATGGGCAGCAGCCATCATCATCATCAGCAGCGGCTTGGTCCGCGCGGACGCCAT	60		
QY	61	ATGGGCTAGCATGACTGGTGGACAGCAAAATGGTCCGGGATCCAGCGCCTCTGTTTTGATG	120		
Db	61	ATGGGCTAGCATGACTGGTGGACAGCAAAATGGTTCGGATCCCCCATGGCGAGTCTTCCCG	120		
QY	121	GCTAAACCCCAAGTGACTATCTTAGCCACAGGAGGCCACCATGCG	164		
Db	121	GAGATCCCGAAGGTGCAGTTCCGAGGGCAAGAGATCCACCAACCC	164		

Search completed: August 27, 2003, 11:17:57  
Job time : 847.661 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:35:29 ; Search time 3475.96 Seconds  
(without alignments)  
7922.115 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133

Sequence: 1 atgggcagcagccatcatca.....aagaaagggaattctttcac 1133

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: \*

1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estmu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_htc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_htc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: em\_gss\_hum: \*  
18: em\_gss\_inv: \*  
19: em\_gss\_pin: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_fun: \*  
22: em\_gss\_mam: \*  
23: em\_gss\_mus: \*  
24: em\_gss\_pro: \*  
25: em\_gss\_rod: \*  
26: em\_gss\_pig: \*  
27: em\_gss\_vrl: \*  
28: gb\_gss1: \*  
29: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	155.2	13.7	983	29	BZ568327
C 2	136	12.0	724	28	AQ990947
C 3	116.8	10.3	766	28	AQ500605
C 4	110.8	9.8	757	28	BH376509

C 5	108.6	9.6	843	28	AQ500657
C 6	106.8	9.4	768	29	BZ570599
C 7	97.8	8.6	553	28	AQ872693
C 8	97.8	8.6	553	28	AQ503225
C 9	96.2	8.5	596	28	AQ501032
C 10	96.2	8.5	630	28	AQ873860
C 11	94.4	8.3	453	28	AQ501722
C 12	94.4	8.3	459	28	AQ501723
C 13	93	8.2	561	28	AQ503362
C 14	89.6	7.9	663	28	BH371112
C 15	87	7.7	560	28	AQ873892
C 16	79.2	7.0	840	29	BZ571432
C 17	75.6	6.7	644	28	AQ157105
C 18	74	6.5	1378	28	BH770857
C 19	72.4	6.4	802	29	BZ571326
C 20	63.8	5.6	545	28	AQ873272
C 21	60.6	5.3	781	29	CNS06FWL
C 22	60.6	5.3	871	29	BZ571362
C 23	55	4.9	427	28	AQ991287
C 24	48	4.2	323	28	AZ049501
C 25	46.8	4.1	455	28	AZ049638
C 26	42.6	3.8	462	28	AQ502215
C 27	42.6	3.8	559	10	BG672925
C 28	41.8	3.7	388	12	BJ363543
C 29	41.8	3.7	492	12	BJ390828
C 30	41.2	3.6	260	12	BJ390253
C 31	41.2	3.6	405	12	BJ358893
C 32	41.2	3.6	473	12	BJ411665
C 33	41.2	3.6	516	12	BJ386769
C 34	41.2	3.6	554	12	BJ364097
C 35	41.2	3.6	554	12	BJ417890
C 36	41.2	3.6	559	12	BJ389977
C 37	41.2	3.6	574	12	BJ336311
C 38	41.2	3.6	588	12	BJ386962
C 39	41.2	3.6	594	12	BJ326765
C 40	41.2	3.6	605	12	BJ386822
C 41	41.2	3.6	619	12	BJ358849
C 42	41.2	3.6	622	12	BJ410498
C 43	41.2	3.6	623	12	BJ325097
C 44	41.2	3.6	623	12	BJ327615
C 45	41.2	3.6	626	12	BJ387400

#### ALIGNMENTS

RESULT 1  
BZ568327/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BZ568327 983 bp DNA linear GSS 17-DEC-2002  
pac2-164\_7476.y2 pac2-164 Pseudomonas aeruginosa genomic clone  
pac2-164\_7476, genomic survey sequence.

BZ568327 GI:27201144

GSS.

Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Pseudomonas.

1 (bases 1 to 983)

Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,

Burns, J.L., Kaul, R. and Olsen, M.V.

Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol., (2002) In press

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

Location/Qualifiers



148 TTTTGACAAACGAAATGCACTCTTTAGATACATTTAGAGCGGATGACAGGGAT 89

662 AAATCGGCACAGTCTTATTTATGCAAGTCGAGTAT 697

88 ATTTAGGTACTTTTCAAAATGATGACGTGGAGTTT 53

RESULT 4

BH376509/c

LOCUS

DEFINITION

AG-ND-137P15.TF.1 ND-TAM Anopheles gambiae genomic clone

AG-ND-137P15, genomic survey sequence.

ACCESSION

BH376509

VERSION

BH376509.1 GI:173222651

KEYWORDS

GSS.

SOURCE

Anopheles gambiae (African malaria mosquito)

ORGANISM

Anopheles gambiae

REFERENCE

1 (bases 1 to 757)

AUTHORS

Hong, Y.S., Hogan, J.R., Wang, X., Sarkar, A., Sim, C., Loftus, B.J., Ren, C., Huff, E.R., Carlile, J.L., Black, K., Zhang, H.-B., Gardner, M.J., and Collins, F.H.

TITLE

Construction of a BAC library and generation of BAC end sequence-tagged connectors for genome sequencing of the African malaria mosquito Anopheles gambiae

JOURNAL

Mol. Genet. Genomics 268 (6), 720-728 (2003)

MEDLINE

22542063

PUBMED

12653398

COMMENT

Other\_GSSs: AG-ND-137P15.TR.1

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjoftus@tigr.org

This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.

Seq primer: M13 For

Class: BAC ends.

Location/Qualifiers

1..757

/organism="Anopheles gambiae"

/mol\_type="genomic DNA"

/strain="PEST"

/db\_xref="taxon:7165"

/clone="AG-ND-137P15"

/clone\_lib="ND-TAM"

/note="Vector: pECBAC1; Site\_1: HindIII"

BASE COUNT

182 a 196 c 192 g 187 t

ORIGIN

Query Match 9.8%; Score 110.8; DB 28; Length 757;

Best Local Similarity 50.4%; Pred. No. 1.8e-21;

Matches 271; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

QY 204 TGGAGCAGTCACCGTTGATAAGCTTCTTGCAGCGCTCCCTGCCATCAACGACCTAGCCAC 263

Db 643 TGCAGCCTCTCTGCGGATGATCTCTTGGCTACGGTCTGATGTAGGCAAGTACGCCAA 584

QY 264 CATCAAGGTGACAGATCTCAAGCATTTGGCTCCCAAGAGATCACGGGTAAGGTGGCT 323

Db 583 GATCCAGGTAAACAACTTATCCAACTTCCATCGGACTACATCGGCGCTGAGCGCTGGT 524

Saccharomycetales; Saccharomycetaceae; Saccharomyces.

1 (bases 1 to 766)

Rose-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., deStages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S. and Snyder, M.

Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption

Unpublished

Contact: Kumar A

Michael Snyder, Dept. of Mol. Cell. and Dev. Biology

Yale University

P.O. Box 208103, New Haven, CT 06520-8103, USA

Tel: 203 432 9949

Fax: 203 432 6161

Email: anuj.kumar@yale.edu

te of mTn-3xHA/lacZ insertion.

Seq primer: GGCTCTCTCTTTGGAGTAC

Class: transposon-tagged.

Location/Qualifiers

1..766

/organism="Saccharomyces cerevisiae"

/mol\_type="genomic DNA"

/db\_xref="taxon:4932"

/lab\_host="E. coli"

/clone\_lib="mTn-3xHA/lacZ Insertion Library"

/note="vector: pHS6-Sal; A yeast genomic DNA library (lacking mitochondrial DNA) was prepared in pHS6-Sal, genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT

223 a 175 c 168 g 198 t

ORIGIN

Query Match 10.3%; Score 116.8; DB 28; Length 766;

Best Local Similarity 50.2%; Pred. No. 2.9e-23;

Matches 289; Conservative 0; Mismatches 287; Indels 0; Gaps 0;

QY 122 CTAAACCCCAAGTCACTATCTAGCCACAGAGGACCATCGTGTGGGGAACTA 181

Db 628 CTTTGGCCATCAATCAAAATTTTGGCACCGCGGTACTATCGCTTCCAAAGTTCCGACAA 569

QY 182 CGCTCAAGAGTAGTACTCTGCTGGAGCAGTCACGTTGATAAGCTTCTTCGAGCGCTCC 241

Db 568 GTGCTACAAACCGCGGTATAGCGGATTAACCGTATATGATTTATAGAGCGCTCC 509

QY 242 CTGCCATCAACGCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAG 301

Db 508 CATCTTTAGTGAGAGGCGAGATCTGGACTATCTTCAAGTGTCTAACTGTGTTTCCACTT 449

QY 302 AGATGACGGTAAAGTGTGGCTTAACTAGCCAGCGTGTCAATGAGCTCTTCGCCCAAA 361

Db 448 CTTTAAACTATACGATCTGATCCCATTTGTATCAGGATCTCCGAGGACATAGCTCTG 389

QY 362 AAGACCGAAGCGGTGATCATCACCATGGAATGACACCATGGAAGAGACCGCTTTCT 421

Db 388 ATGATTACCTGCTGCGGTGTGATCTCATGGAACCGACACTATGAGGAGACAGCTTTCT 329

QY 422 TCCTCAACCTCAGCGTGAAGAACCAAAACCTGTCTGCTTGTAGCGGCATTCGCTCCAG 481

Db 328 TCTTAGATTGACATATAATTCAGAGAGCCAGTATGATCGAGCGCTATCGGTCAG 269

QY 482 GCTCTTCATGATGTGATGCGCCCATGATCTCTATAAGCGCGGTGAATGTAGCGATCA 541

Db 268 CCACTGCCACGCTGCTGTATGGCCCAATGAATTTATATCAGAGTGTCTATCGCTGCTT 209

QY 542 ACAAGCCTCTACTAAACAAAGAGTGTGATTTGTGATGAACGATGAGATTACGCCCA 601

Db 208 CTGAGAAATCACTGGGTGCTGGGACGATGATCACTCTAAACGATCTATTTGCTCTGGT 149

QY 602 GAGAACGACCAAGCTCAACACCAACCGCAGTCAATGCTTTGCTTCGCCCAACACAGGTA 661

QY 324 TAAACTAGCCCAAGCGTGTCAATGAGCTCTCGCCCAAAAGAGACCGGCGTGATCAT 383  
 Db 523 CCAGCTCACTGCTGCTGTGCAAGCGTATCAACAGTCAGAGCTTCTGGCGTGATCGT 464  
 QY 384 CACCCATGGAACCTGACACCAATGGAAGAGACCGCTTTCTCTCAACCTCAGCGTGAAAG 443  
 Db 463 TTCTCATGCGACAGATACGCTGGAGGAAACAGCTTTTGGCTTGAATCTGACATGAAGTC 404  
 QY 444 CCAAAAACCTGTCTGCGCTTGTAGCGCCATGCGTCCAGGCTCTTCATGAGTCTGATGG 503  
 Db 403 CATTAACCAATCGTCTGATTTGGCGCAGACGCAACGCGCTCTTCAGACTTTGACGG 344  
 QY 504 CCCATGAATCTCTATAAGCCCGTGAATGTAGCGATCAACAAAGCCTCTACTAACAAAG 563  
 Db 343 TCGCGCAACTTGTCTTAACGCTGTGCGCATCGTGTGTATCCCAATCCAAGATCGTGG 284  
 QY 564 AGTGGTGATTTGATCAAGATGAGATTACGCGCGCAGAGAGCAACCAAGCTCAACAC 623  
 Db 283 CGCATGCTGGCGATGAACAACAGATCAACGCTACCAAGTATGTCAACAGACACAC 224  
 QY 624 CACCGCAGTCAATGCTTTGCTTTCGCCCAACACACAGGTAAATCGGCATTATGATGG 683  
 Db 223 TGCCCAAGCTTGAACCTTTAAGCTGGTGAGTTCGGCTTATTTGGTGGAGTCTACCCAGA 164  
 QY 684 CAAAGTCGAGTATTTCACTCAATCCGCTTCGACCTCACACCTTGCAAGTGAGTTTGAT 741  
 Db 163 CAAAGTCGTTTACCGCACTAGCGCGATCGGCAATTTAAACTTTCTATCAAGGCAGAT 106

## RESULT 5

AQ500657/c

## LOCUS

V34rII mTn-3XHA/lacZ Insertion Library Saccharomyces cerevisiae  
 genomic 5', genomic survey sequence.

AQ500657

AQ500657.1

GI:4706367

GSS.

Saccharomyces cerevisiae (baker's yeast)

Saccharomyces cerevisiae

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

1 (bases 1 to 843)

Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,

desEtages, S.A., Cheung, K.-H., Sheehan, A., Symonitis, D., Jansen, R.,

Unesky, L., Heideman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,

Hager, K., Miller, P., Koeder, G.S. and Snyder, M.

Large-Scale Analysis of the Yeast Genome by Transposon Tagging and

Gene Disruption

Unpublished

Contact: Kumar A

Michael Snyder, Dept. of Mol. Cell. and Dev. Biology

Yale University

P.O. Box 208103, New Haven, CT 06520-8103, USA

Tel: 203 432 9949

Fax: 203 432 6161

Email: anuj.kumar@yale.edu

te of mTn-3XHA/lacZ insertion.

Seq primer: GGCCTTCTTCTTGGAGTAC

Class: transposon-tagged.

Location/Qualifiers

1..843

/organism="Saccharomyces cerevisiae"

/mol\_type="genomic DNA"

/db\_xref="taxon:4932"

/lab\_host="E. coli"

/clone\_lib="mTn-3XHA/lacZ Insertion Library"

/notes="Vector: pHS6-Sal; A yeast genomic DNA library

(lacking mitochondrial DNA) was prepared in pHS6-Sal;

genomic DNA was size-fractionated (DNA of roughly 2-3 kb

in length) prior to cloning. This library was

subsequently mutagenized with a mTn-3XHA/lacZ

minitransposon containing lacZ, URA3, and tet resistance."

245 a 182 c 182 g 224 t 10 others

( BASE COUNT

## FEATURES

source

## ORIGIN

Query Match 9.6%; Score 108.6; DB 28; Length 843;  
 Best Local Similarity 50.7%; Pred. No. 8.6e-21;  
 Matches 255; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 208 GCAGTCACCGTTGATAAGCTTCTTGCAGCGGTCCCTGCCATCAACGACCTAGCCACCATC 267  
 Db 679 GGATTAACCGTAAATGATNTAATAGAGCGGTCATCTTTAGCTGAGAGGACATCTG 620  
 QY 268 AAGGTGAACAGATCTCAAGCATTTGGTCCCAGAGATACGGGTAAAGTGTGGCTTAA 327  
 Db 619 GACTATCTTCAAGTGTCTAAACGTTGGTTCCANATCTTTANACTATATACGATCTGATCCCA 560  
 QY 328 CTAGCCAAAGCGTGTCAATGAGCTCTCGCCCAAAAGAGACCGGTCGATGATGATGAT 387  
 Db 559 TTGATACCGGTATCTCCGAGCATTAGCTTCTGATGACTACGCTGGTGGTGTGCACT 500  
 QY 388 CATGGAATGACACCATGGAAGAGACCGCTTTCTTCTCAACCTCAACGCTGAAAGCCCAA 447  
 Db 499 CATNGACCGACACTATGAGGAGAGACAGCTTTCTTCTTAGATTTGCCATAAATTCAGAG 440  
 QY 448 AAACCTGTCTGCTTGTAGCGCCCATGCGTCCAGCTCTTCCATGAGTGTGATGCGCCC 507  
 Db 439 AAGCAGATATGATCGCAGCGCTATGCGTCCAGCCTGCGCACCTGCTGCTGATGGCCCA 380  
 QY 508 ATGAATCTTATAACCGCGTGAATGTAGCGATCAACAAAGCTCTTACTAACAAAGAGTG 567  
 Db 379 ATGAATTTATATCAAGCAGTGTCTATTTGCTCTTGAGAATCACTGGGTGCTGGCAG 320  
 QY 568 GTGATTGTGATGAACGATGAGATTCACGCGCCAGAGAGCGACCAAGCTCAACACCAACC 627  
 Db 319 ATGACTACTCTAAACGATCGTATTTGCTCTGGGTTTGGACAACGAAATGAATGCCAAC 260  
 QY 628 GCAGTCAATGATTTGCTTTCGCCCAACACACAGGTAAATCGGCACAGTCTATTATGGCAA 687  
 Db 259 TCTTTAGATACATTACAGCGGATGAACAGGATATTAGGTACTTTTCAATGATGAC 200  
 QY 688 GTCAGATTTTCACTCAATCCGT 710  
 Db 199 GTGAGATTTTACTACCACCACT 177

## RESULT 6

BZ570599/c

## LOCUS

BZ570599 768 bp DNA linear GSS 17-DEC-2002  
 msh2\_1446.y2 msh Pseudomonas aeruginosa genomic clone msh2\_1446,  
 genomic survey sequence.

BZ570599

BZ570599.1

GI:27205660

GSS.

Pseudomonas aeruginosa

Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Pseudomonas.

1 (bases 1 to 768)

Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,

Burns, J.L., Kaul, R. and Olsen, M.V.

Whole-Genome-Sequencing variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol., (2002) In press

Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

Location/Qualifiers

1..768

/organism="Pseudomonas aeruginosa"

/mol\_type="genomic DNA"

source

/strain="MSH"  
/db\_xref="taxon:287"  
/clone="msh2\_1446"  
/note="Environmental isolate. Whole genomic shotgun library."  
library."

```

BASE COUNT      116 a   246 c   239 g   165 t   2 others
ORIGIN

Query Match      9.4%; Score 106.8; DB 29; Length 768;
Best Local Similarity 53.6%; Pred. No. 2.8e-20;
Matches 287; Conservative 0; Mismatches 243; Indels 5; Gaps 3;

QY 331 GCCAAGCGTGTCAATGAGCTCTCGCCCAAAAGAGACCGAGCGTGTATCATACCCCAT 390
Db 610 GGCATATACCGTGCCTAGCTGCGGATACGGAACAGCTGCGGATCATCGTATACCCCC 551
QY 391 GGAACCTGACACCATGGAAGAGACCGCTTTCTTCTCAACCTCACGCTGAAAGCCAAAAA 450
Db 550 GGTACCGACACCT-GAAGAGACCGCCACTTCTTGACCTGTGTGAGCACACCGGAGAG 492
QY 451 CTTGTCTGCTTGTAGGCGCATGCTCCAGCTCTTCCATGAGTGTCTGATGCCCCCATG 510
Db 491 CCTATCGTGTGTGCTGCTCGATGCGCCCGGCACCGCNATGTCCGCGACGCGCATGCTC 432
QY 511 AATCTCTATAACCGCGTGAATGTAGCGATCAACAAAGCCTCTACTAACAAAGAGGTGTG 570
Db 431 AACCTGTACACGCGG-GCGGTGCGCGGACAGTCCGACCGCGGCAAGGCGGTGCTG 373
QY 571 ATTGTGATGAACGATGAGATTCAACGCGCGCAGAGACGACCAAGCTCAACACCCGCA 630
Db 372 ATCACCATGAACGAGAGATCTCTCCGCGCGCAGCGAGCAAGATGCTCAACATCAAG 313
QY 631 GTCATGCAATTGCTTCGCGCAACAGACAGTAAATCGGCACAGTCTATTATGGCAAGTC 590
Db 312 ACCGAAGCGTT---CAAGAGCGCGTGGGCGCCGCTGGCGATGCTGTGAGGCGCAAGAG 256
QY 691 GAGTATTTCACTCAATTCGCTGACCTCACACCTTGCAGTGTGAGTTTCATATTAGCAA 750
Db 255 TACTGTGTCGCGCACCGTGAAGCGGACACGGTGAATCTGAGTTCGATCAATCAAGCAG 196
QY 751 ATCGAAGAACTCCCGAGTGCATATTCTTTAGCTCACCCGATGATGATGATGATTTTA 810
Db 195 ATCTCGCGCTGCTCGCTCGGTGGAATCGCTACAGCTACGGAACGCTCAGCGACACCGC 136
QY 811 GTCATGAGCCCTTCAGCAGGAGCAAGGATCATCATCGCATGCGCATGGCA 865
Db 135 TACAAGGCGCTGCGCCAGCGCGGCGCAAGGCGATCATCATGCGCGCAGGGGGA 81

```

```

RESULT 7
AQ872693/c
LOCUS      553 bp   DNA      linear      GSS 08-NOV-1999
DEFINITION V34E5 mTn-3xHA/lacZ Insertion Library, strain AB972 Saccharomyces cerevisiae genomic 5', genomic survey sequence.
ACCESSION  AQ872693
VERSION     AQ872693.1 GI:6285000
KEYWORDS   GSS.
SOURCE     Saccharomyces cerevisiae (baker's yeast)
ORGANISM   Saccharomyces cerevisiae
REFERENCE  1 (bases 1 to 553)
AUTHORS   Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
TITLE     Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption
JOURNAL   Unpublished
COMMENT   Contact: Kumar A
          Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
          Yale University

```

P.O. Box 208103, New Haven, CT 06520-8103, USA  
Tel: 203 432 9949  
Fax: 203 432 6161  
Email: anuj.kumar@yale.edu  
te of mTn-3xHA/lacZ insertion.  
Seq primer: GGCCTTCTTTTGTGGAAGTAC  
Class: transposon-tagged.  
Location/Qualifiers

FEATURES  
source

1..553  
/organism="Saccharomyces cerevisiae"  
/mol\_type="genomic DNA"  
/strain="AB972 - trp1 r(0) (S288C background)"  
/db\_xref="taxon:4932"  
/lab\_host="E. coli"  
/clone\_lib="mTn-3xHA/lacZ Insertion Library, strain AB972"  
/note="Vector: pHS6-Sal; A yeast genomic DNA library was prepared in pHS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mTn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 152 a 127 c 124 g 150 t

```

ORIGIN

Query Match      8.6%; Score 97.8; DB 28; Length 553;
Best Local Similarity 53.9%; Pred. No. 1.1e-17;
Matches 201; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 338 GTGTCAATGAGTCTCGCCCAAAAGAGACCGAGCGTGTATCATCACCATGGAATG 397
Db 548 GTATCTCCGAGCAGCTAGCTTCTGATGACTACGCTGTGCTGCTGATGAGCCG 489
QY 398 ACACCATGAGAGAGACCGCTTTCTTCTCAACTCACGCTGAAAGCCAAACACCTGTCT 457
Db 488 ACATATGAGAGAGACAGCTTTCTTCTAGATTGACCAATAATTGAGAGACCATAT 429
QY 458 GCCTTTGAGGCGCCATGCTCCAGGCTTTCCATGAGTGTCTGATGCCCATGAATCTCT 517
Db 428 GTATCGAGCGCTATGCTCCAGCAGCTGCCAGCTGCTGATGCCCAATTAAT 369
QY 518 ATACGCGCTGATGTAGCGATCAACAAAGCTTCTACTAACAAAGAGTGTGATGTGA 577
Db 368 ATCAAGCAGTGTATTGCTGCTTCTGAGAAATCACTGGGTGCTGCGATGATCACTC 309
QY 578 TGAACGATGAGATTCAACGCGCGCAGAGAGACCAAGCTCAACACCCGCGAGTCAATG 637
Db 308 TAAACGATGATTTGCTCTGGGTTTGGACAAAGAAATGATGCCACTCTTTAGATA 249
QY 638 CATTTGCTTTCGCCCAACACAGGTAAATCGGCACAGTCTATTATGCGAAGTCGATTT 697
Db 248 CATTCAGAGCGGATGAACAGGGATATTAGGTACTTTTCAATGATGACGTGGAGTTT 189
QY 698 TCACTCAATCCGT 710
Db 188 ACTACCCACAGT 176

```

```

RESULT 8
AQ503225/c
LOCUS      553 bp   DNA      linear      GSS 29-APR-1999
DEFINITION V34E05 mTn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae genomic 5', genomic survey sequence.
ACCESSION  AQ503225
VERSION     AQ503225.1 GI:4705771
KEYWORDS   GSS.
SOURCE     Saccharomyces cerevisiae (baker's yeast)
ORGANISM   Saccharomyces cerevisiae
REFERENCE  1 (bases 1 to 553)
AUTHORS   Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A., desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R., Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,

```

TITLE Hager, K., Miller, P., Roeder, G.S. and Snyder, M.  
 JOURNAL Large-Scale Analysis of the Yeast Genome by Transposon Tagging and  
 COMMENT Gene Disruption  
 Unpublished  
 Contact: Kumar A  
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology  
 Yale University  
 P.O. Box 208103, New Haven, CT 06520-8103, USA  
 Tel: 203 432 9949  
 Fax: 203 432 6161  
 Email: anuj.kumar@yale.edu  
 te of mtn-3xHA/lacZ insertion.  
 Seq primer: GGCCTCTCTTTTGGAGTAC  
 Class: transposon-tagged.  
 Location/Qualifiers

## FEATURES

source  
 1. .553  
 /organism="Saccharomyces cerevisiae"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4932"  
 /lab\_host="E. coli"  
 /clone\_lib="mtn-3xHA/lacZ Insertion Library"  
 /note="Vector: pHS6-Sal; A yeast genomic DNA library  
 (lacking mitochondrial DNA) was prepared in pHS6-Sal;  
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb  
 in length) prior to cloning. This library was  
 subsequently mutagenized with a mtn-3xHA/lacZ  
 minitransposon containing lacZ, URA3, and tet resistance."  
 BASE COUNT 152 a 127 c 124 g 150 t  
 ORIGIN

Query Match 8.6%; Score 97.8; DB 28; Length 553;  
 Best Local Similarity 53.9%; Pred. No. 1.1e-17;  
 Matches 201; Conservative 0; Mismatches 172; Indels 0; Gaps 0;  
 QY 338 GTGTCATGAGCTCTCGCCCAAAAGAGACCGGCGTGCATCATCAACCCATGGAATG 397  
 Db 548 GTATCTCCGAGGCACTAGCTTCTGATGACTACGCTGTGGCGGTGTCACTCATGGACCG 489  
 QY 398 ACACCATGGAGAGACCGCTTCTCTCAACCTCAGCGTGAAAGCCAAAACCTGCT 457  
 Db 488 ACATATGGAGGAGACACTTCTCTTAGATTGACATAAATTCAGAGACGATAT 429  
 QY 458 GCCTTGTAGCGGCATCGCTCCAGGCTCTTCCATGAGTGTGATGGCCCCATGAATCTCT 517  
 Db 428 GTATCGCAGCGCTATCGCTCCAGCCTGCGACGCTGCTGTGATGGCCCAATGAATTAT 369  
 QY 518 ATACGCGCGTGAATGACGATCAACAAGCCTCTACTACAAAGGAGTGTGATTGTA 577  
 Db 368 ATCAAGCAGTGTCTATTGCTGCTTCTGAGAAATCACTGGGTCGCGCACGATGATCACTC 309  
 QY 578 TGAACGATGAGATTACGCGCCGACGAGAGCGACCAAGCTCAACACCAACCGCAGTCAATG 637  
 Db 308 TAAACGATCGATTGCTCTGGGTTTGGACACAGAAATGAATGCCACTTTAGATA 249  
 QY 638 CATTTGTTCCGCCAACACAGGTAAATCGGCACAGCTCTATTATGGCAAAAGTCGAGTATT 697  
 Db 248 CATTCAGACGGATGAACAGGGATATTAGGTACTTTTCAAAATGATGACGTGGAGTTT 189  
 QY 698 TCACCTCAATCCGT 710  
 Db 188 ACTACCACCACT 176

RESULT 9  
 AQ501032/c  
 LOCUS V28H11 mtn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae  
 DEFINITION genomic 5', genomic survey sequence.  
 ACCESSION AQ501032  
 VERSION AQ501032.1  
 KEYWORDS GSS.  
 SOURCE Saccharomyces cerevisiae (baker's yeast)  
 ORGANISM Saccharomyces cerevisiae

REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 1 (bases 1 to 596)  
 AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,  
 desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,  
 Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,  
 Hager, K., Miller, P., Roeder, G.S. and Snyder, M.  
 TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and  
 JOURNAL Gene Disruption  
 COMMENT Unpublished  
 Contact: Kumar A  
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology  
 Yale University  
 P.O. Box 208103, New Haven, CT 06520-8103, USA  
 Tel: 203 432 9949  
 Fax: 203 432 6161  
 Email: anuj.kumar@yale.edu  
 te of mtn-3xHA/lacZ insertion.  
 Seq primer: GGCCTCTCTTTTGGAGTAC  
 Class: transposon-tagged.  
 Location/Qualifiers

## FEATURES

source  
 1. .596  
 /organism="Saccharomyces cerevisiae"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4932"  
 /lab\_host="E. coli"  
 /clone\_lib="mtn-3xHA/lacZ Insertion Library"  
 /note="Vector: pHS6-Sal; A yeast genomic DNA library  
 (lacking mitochondrial DNA) was prepared in pHS6-Sal;  
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb  
 in length) prior to cloning. This library was  
 subsequently mutagenized with a mtn-3xHA/lacZ  
 minitransposon containing lacZ, URA3, and tet resistance."  
 BASE COUNT 171 a 132 c 130 g 163 t  
 ORIGIN

Query Match 8.5%; Score 96.2; DB 28; Length 596;  
 Best Local Similarity 54.0%; Pred. No. 3.4e-17;  
 Matches 197; Conservative 0; Mismatches 168; Indels 0; Gaps 0;  
 QY 346 GAGCTCTCGCCCAAAAGAGACCGGCGTGCATCATCAACCCATGGAATGACCATG 405  
 Db 533 GAGGCACCTAGCTTCTGTGATGATTACGCTGTGGCGGTGTCACTCATGGACCGACATG 474  
 QY 406 GAAGAGACCGCTTCTCTCAACCTCAGCGTGAAAGCCAAAACCTGCTGCTGTGTA 465  
 Db 473 GAGGAGACGCTTCTCTTAGATTGACCAATAAATTCAGAAAGCAGTATGATCGCA 414  
 QY 466 GCGCCATCGCTCCAGGCTCTTCCATGATGCTGATGGCCCATGAATCTCTATAACGCC 525  
 Db 413 GCGCTATCGTCCAGCCACCTGCCACGCTGCTGATGGCCCAATGAATTTATATCAAGCA 354  
 QY 526 GTGAATGTAGCGATCAAAAGCCTCTACTAAAGAGAGTGTGTGATGATGAACGAT 585  
 Db 353 GTGCTATTGTGCTTCTTGAGAAATCACTGGGTCGTGGCAGCATGATCACTCTTAAACGAT 294  
 QY 586 GAGATTTCACGCGCCGACGAGAGCGACCAAGCTCAACACCCAGCAGTCAATGATTTGCT 645  
 Db 293 CGTATGCTCTGGGTTTGGACAAAGAAATGAATGCCAATCTTTTAGATACATTCAGA 234  
 QY 646 TCGCCCCAACACAGGTAAATCGGCACAGTCTATTATGGCAAAAGTCGAGTATTTCACTCAA 705  
 Db 233 CGCGATGAACAGGGATATTTAGTTACTTTTCAATGATGACGTGGAGTTTACTACCCA 174  
 QY 706 TCCGT 710  
 Db 173 CCAGT 169

RESULT 10  
 AQ873860/c  
 LOCUS V82E1 mtn-3xHA/lacZ Insertion Library, strain AB972 Saccharomyces  
 DEFINITION

cerevisiae genomic 5', genomic survey sequence.

ACCESSION AQ873860  
 VERSION AQ873860.1 GI:6286104  
 KEYWORDS GSS.  
 SOURCE Saccharomyces cerevisiae (baker's yeast)  
 ORGANISM Saccharomyces cerevisiae  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes;  
 1 (bases 1 to 630)  
 Ros-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,  
 DesEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,  
 Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,  
 Hager, K., Miller, P., Roeder, G.S. and Snyder, M.  
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and  
 Gene Disruption  
 Unpublished  
 Contact: Kumar A  
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology  
 Yale University  
 P.O. Box 208103, New Haven, CT 06520-8103, USA  
 Tel: 203 432 9949  
 Fax: 203 432 6161  
 Email: anuj.kumar@yale.edu  
 te of mTn-3xHA/lacZ insertion.  
 Seq primer: GGCTTCTTCTTGTGAAGTAC  
 Class: transposon-tagged.  
 Location/Qualifiers  
 1..630  
 /organism="Saccharomyces cerevisiae"  
 /mol\_type="genomic DNA"  
 /strain="AB972 - trp1 x(0) (S288C background)"  
 /db\_xref="taxon:4932"  
 /lab\_host="E. coli"  
 /clone\_lib="mTn-3xHA/lacZ Insertion Library, strain AB972"  
 /note="Vector: pHS6-Sal; A yeast genomic DNA library was  
 prepared in pHS6-Sal; genomic DNA was size-fractionated  
 (DNA of roughly 2-3 kb in length) prior to cloning. This  
 library was subsequently mutagenized with a mTn-3xHA/lacZ  
 minitransposon containing lacZ, URA3, and tet resistance."

## FEATURES

source

BASE COUNT 179 a 138 c 138 g 175 t

## ORIGIN

Query Match 8.5%; Score 96.2; DB 28; Length 630;  
 Best Local Similarity 53.6%; Pred. No. 3.5e-17;  
 Matches 200; Conservative 0; Mismatches 173; Indels 0; Gaps 0;  
 QY 338 GTGTCAATGAGTCTTCGCCCAAAAGAGACCGAGCGGTGATCATCACCCTGGAACGTG 397  
 DB 550 GTATCTCCGAGGCACTAGCTTCTGTATGACTACGCTGGTGTGTTCACTCATGCGACG 491  
 QY 398 ACACCATGGAAGAGACCGCTTCTTCTCACTCAGGTGAAAGCCCAAAACCTGTCT 457  
 DB 490 ACACATGAGGAGAGACGATATCTTCTTAGATTGACCAATAATTCAGAGAAGCCAGTAT 431  
 QY 458 GCCTTTGAGCGGCATCGCTCCAGGCTCTTCCATGAGTGTCTGATGCCCCCATGAATCTCT 517  
 DB 430 GTATCCAGCGCGTATGCGTCCAGCCACTGCCACGCTGCTGTATGCCCAATGATTTAT 371  
 QY 518 ATAACCCCGTGAATGATGAGGATCAACAAGCCCTCTACTACAAGAGATGGTGTGATGTA 577  
 DB 370 ATCAAGCAGTGTCTATTGCTGCTTCTGAGAAATCACTGGGTCTGTCGACGATGATCACTC 311  
 QY 578 TGAACCATGATGATTCAGCGCCGACAGAGAGCGACCAAGCTCAACACCGGAGTCAATG 637  
 DB 310 TAAACGATCGTATTGCTCTGGGTTTGGACAAAGCAATGAAATGCAATCTTTAGATA 251  
 QY 638 CATTTGCTCGCCCAACACAGGTAATAATCGGCACAGTCTATTATGGCAAGTGCAGTATT 697  
 DB 250 CATTCAGCGGATGACAGGGATATTAGTTACTTTTCAATGATGACGTGAGTTT 191  
 QY 698 TCACTCAATCCGT 710

Db 190 ACTACCCACCAGT 178

## RESULT 11

AQ501722/c

LOCUS

DEFINITION

genomic 5', genomic survey sequence.

ACCESSION AQ501722

VERSION AQ501722.1 GI:4707372

KEYWORDS GSS.

SOURCE Saccharomyces cerevisiae (baker's yeast)

ORGANISM Saccharomycetes cerevisiae

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes;  
 1 (bases 1 to 453)  
 Ros-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,  
 DesEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,  
 Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,  
 Hager, K., Miller, P., Roeder, G.S. and Snyder, M.  
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and  
 Gene Disruption  
 Unpublished  
 Contact: Kumar A  
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology  
 Yale University  
 P.O. Box 208103, New Haven, CT 06520-8103, USA  
 Tel: 203 432 9949  
 Fax: 203 432 6161  
 Email: anuj.kumar@yale.edu  
 te of mTn-3xHA/lacZ insertion.  
 Seq primer: GGCTTCTTCTTGTGAAGTAC  
 Class: transposon-tagged.  
 Location/Qualifiers  
 1..453  
 /organism="Saccharomyces cerevisiae"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4932"  
 /lab\_host="E. coli"  
 /clone\_lib="mTn-3xHA/lacZ Insertion Library"  
 /note="Vector: pHS6-Sal; A yeast genomic DNA library  
 (lacking mitochondrial DNA) was prepared in pHS6-Sal;  
 genomic DNA was size-fractionated (DNA of roughly 2-3 kb  
 in length) prior to cloning. This library was  
 subsequently mutagenized with a mTn-3xHA/lacZ  
 minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 111 a 121 c 115 g 105 t

## ORIGIN

Query Match 8.3%; Score 94.4; DB 28; Length 453;  
 Best Local Similarity 56.9%; Pred. No. 9.8e-17;  
 Matches 173; Conservative 0; Mismatches 131; Indels 0; Gaps 0;  
 QY 338 GTGTCAATGAGTCTTCGCCCAAAAGAGACCGAGCGGTGATCATCACCCTGGAACGTG 397  
 DB 372 GTATCTCCGAGGCACTAGCTTCTGTATGACTACGCTGGTGTGTTCACTCATGCGACG 313  
 QY 398 ACACCATGGAAGAGACCGCTTCTTCTCACTCAGGTGAAAGCCCAAAACCTGTCT 457  
 DB 312 ACACATGAGGAGAGACGCTTCTTCTTAGATTGACCAATAATTCAGAGAAGCCAGTAT 253  
 QY 458 GCCTTTGAGCGGCATCGCTCCAGGCTTTCATGAGTGTGATGCCCCCATGAATCTCT 517  
 DB 252 GTATCGAGGCGCTATGCTCCAGGCACTGCCACGCTGCTGTATGCGCCCAATGAATTTAT 193  
 QY 518 ATAACCCCGTGAATGATGAGGATCAACAAGCCCTCTACTACAAGAGTGGTGTGATGTA 577  
 DB 192 ATCAAGCAGTGTCTATTGCTGCTTCTGAGAAATCACTGGGTCTGTCGACGATGATCACTC 133  
 QY 578 TGAACCATGAGATTCACGCGCGCAGAGAGAGCGACCAAGCTCAACACCGGAGTCAATG 637  
 DB 132 TAAACGATCGTATTGCTCTGGGTTTGGACAAAGCAATGAATGCAATCTTTAGATA 73

```

QY 638 CATT 641
Db 72 CATT 69

RESULT 12
AQ501723/c 459 bp DNA linear GSS 29-APR-1999
LOCUS V15G2 mTn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
DEFINITION genomic 5', genomic survey sequence.
ACCESSION AQ501723
VERSION AQ501723.1 GI:4707373
KEYWORDS GSS.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
REFERENCE 1 (bases 1 to 459)
AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
JOURNAL Unpublished
COMMENT Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCCTCTCTTTGGAAGTAC
Class: transposon-tagged.
Location/Qualifiers
FEATURES
source
1..459
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library"
/note="Vector: pHS6-Sal; A yeast genomic DNA library
(lacking mitochondrial DNA) was prepared in pHS6-Sal;
genomic DNA was size-fractionated (DNA of roughly 2-3 kb
in length) prior to cloning. This library was
subsequently mutagenized with a mTn-3xHA/lacZ
minitransposon containing lacZ, URA3, and tet resistance."
BASE COUNT 116 a 119 c 120 g 102 t
ORIGIN
Query Match 8.3%; Score 94.4; DB 28; Length 459;
Best Local Similarity 56.9%; Pred. No. 9.9e-17;
Matches 173; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 338 GTCTCATGAGCTCTCGCCCAAAAGACGAGCGGTGATCATCAACCCATGAACTG 397
Db 349 GTATCTCCGAGGACATAGCTCTGATGACTACGCTGTGCGGTGTCTACTCATGGACCG 290
QY 398 ACACCATGGAAGAGACCGCTTCTCTCAACCTCAGCGTGAAGCCAAACCTGTCT 457
Db 289 ACATATGAGGAGACAGCTTCTCTTAGATTGACCATTAATTCAGAGAAGCCAGTAT 230
QY 458 GCCTTGTAGCGGCATCGCTCCAGGCTCTTCCATGAGTGTGATGGCCCAATGAATCTCT 517
Db 229 GTATCGAGCGGTATGCGTCCAGCCACTGCCAGTCTGCTGATGGCCCAATGAATTTAT 170
QY 518 ATAACCCGCTGATGAGCATCAACAAAGCTCTACTAACAAGAGGAGTGTGATGTA 577
Db 169 ATCAAGCAGTGTCTATTGCTTCTGAGAAATCACTGGGTGCGGACCATGATCACTC 110
QY 578 TGAACGATGAGATTACGCGCCAGAGAGGACCAAGCTCAACACACCGACGCTCAATG 637

RESULT 13
AQ503362/c 561 bp DNA linear GSS 29-APR-1999
LOCUS V57D10 mTn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
DEFINITION genomic 5', genomic survey sequence.
ACCESSION AQ503362
VERSION AQ503362.1 GI:4705908
KEYWORDS GSS.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
REFERENCE 1 (bases 1 to 561)
AUTHORS Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,
Umansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S. and Snyder, M.
TITLE Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
JOURNAL Unpublished
COMMENT Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCCTCTCTTTGGAAGTAC
Class: transposon-tagged.
Location/Qualifiers
FEATURES
source
1..561
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mTn-3xHA/lacZ Insertion Library"
/note="Vector: pHS6-Sal; A yeast genomic DNA library
(lacking mitochondrial DNA) was prepared in pHS6-Sal;
genomic DNA was size-fractionated (DNA of roughly 2-3 kb
in length) prior to cloning. This library was
subsequently mutagenized with a mTn-3xHA/lacZ
minitransposon containing lacZ, URA3, and tet resistance."
BASE COUNT 158 a 128 c 118 g 157 t
ORIGIN
Query Match 8.2%; Score 93; DB 28; Length 561;
Best Local Similarity 53.4%; Pred. No. 3e-16;
Matches 195; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 346 GAGCTCTCGCCCAAAAGACGAGCGGTGATCATCAACCCATGAACTGACACCATG 405
Db 514 GAGGATTAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 455
QY 406 GAAGAGACCGCTTCTTCTCAACCTCAGCGTGAAGCCAAAGCCAAACCTGTCTGCTTGTGA 465
Db 454 GAGGAGACAGCTTCTTCTTAGATTGACCATAGATTACAGAGAGCCAGTATGATCGCA 395
QY 466 GCGCCCATGCGTCCAGGCTCTTCCATGAGTGTGATGGCCCATGATGATGATGATGATGATG 525
Db 394 GCGCTATGCGTCCAGCCACTGCCAGCTGCTGCTGATGGCCCAATGAATTTATATCAAGCA 335
QY 526 GTGAATGTAGCATCAACAAAGCCCTCTACTAACAAGAGGAGTGTGATGATGATGATGATGATG 585
Db 334 GTGTCTATTGATGCTTCTTGAGAAATCACTGGGTGCGGACCATGATGATGATGATGATGATG 275

```

QY 586 GAGATTACGCGCGCAGAGAGCGACCAAGCTCAACACCGCGAGTCAATGCAATTGCT 645  
 |||||  
 Db 274 CGTATTGCTCTGGGTTTGGACACGAAATGAATGCCAACTCTTTAGATACATTGCA 215  
 |||||  
 QY 646 TCGCCCAACACAGCTAAATCGGCACAGTCTATTATGCAAGTGCAGTATTCTACTCA 705  
 |||||  
 Db 214 GCGATGACACAGGATATTTAGGTACTTTTCAATGATGACGTGGAGTTTACTACCCA 155  
 |||||  
 QY 706 TCCGT 710  
 |||||  
 Db 154 CCACT 150  
 |||||

RESULT 14  
 BH371112/c  
 LOCUS  
 DEFINITION AG-ND-137P15.TF ND-TAM Anopheles gambiae genomic clone AG-ND-137P15  
 , genomic survey sequence.

ACCESSION  
 VERSION BH371112.1 GI:17317237  
 KEYWORDS  
 SOURCE  
 ORGANISM Anopheles gambiae (African malaria mosquito)  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;  
 Anopheles.

REFERENCE  
 AUTHORS Hong, Y.S., Hogan, J.R., Wang, X., Sarkar, A., Sim, C., Loftus, B.J., Ren,  
 C., Huff, E.R., Carlile, J.L., Black, K., Zhang, H.-B., Gardner, M.J.  
 and Collins, F.H.

TITLE Construction of a BAC library and generation of BAC end  
 sequence-tagged connectors for genome sequencing of the African  
 malaria mosquito Anopheles gambiae

JOURNAL  
 MEDLINE Mol. Genet. Genomics 268 (6), 720-728 (2003)  
 PUBMED 22542063

COMMENT  
 Other\_GSSs: AG-ND-137P15.TR  
 Contact: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: bjloftus@tigr.org

This clone is from an A. gambiae BAC library (ND-TAM) provided by  
 F.H. Collins and sequenced by The Institute for Genomic Research  
 (TIGR). The BAC library was generated from A. gambiae PEST strain  
 DNA. All DNA was extracted from newly hatched first instar larvae  
 to minimize the inclusion of DNA from microorganisms that inhabit  
 the gut. The DNA is derived from mixed sexes of larvae. The BAC  
 library was constructed at Texas A&M University BAC Center  
 University, College Station, Texas 77843-2123, USA using a HindIII  
 partial digest.

Seq primer: M13 For  
 Class: BAC ends.

FEATURES  
 source  
 Location/Qualifiers  
 1..663  
 /organism="Anopheles gambiae"  
 /mol\_type="genomic DNA"  
 /strain="PEST"  
 /db\_xref="taxon:7165"  
 /clone\_lib="AG-ND-137P15"  
 /notes="Vector: pECBAC1; Site 1: HindIII"  
 BASE COUNT 160 a 175 c 168 g 160 t

Query Match 7.9%; Score 89.6; DB 28; Length 663;  
 Best Local Similarity 50.4%; Pred. No. 3.4e-15;  
 Matches 244; Conservative 0; Mismatches 239; Indels 1; Gaps 1;

QY 259 GCCACATCAAGGTCACAGATCTCAAGCATTTGGCTCCCAAGAGATGACGGTAAGTG 318  
 |||||

Db 589 GCCAAGATCCAGGTAAACAACTTATCCACGTTCCATCGGACTACATGGGCGCTCAGCGC 530  
 QY 319 TGGCTTAACTAGCCAAAGCGGTGCAATGAGCTCTCGCCCAAAGAGACCGAAGCCGTG 378  
 |||||  
 Db 529 TGGGTCCAGCTCACTGCTGTGTGCAAGACGCTATCAACAAGTCAGACGTTTGTGGCGTG 470  
 |||||  
 QY 379 ATCATCACCACATGGAATGACACCATGGAAGAGACCGCTTTCTTCTCAACCTCAGCGTG 438  
 |||||  
 Db 469 ATCGTTTCTCATGGCACAGATACGCTGGAGAAACAGCTTTTGGCTTGTATCTGACAGTG 410  
 |||||  
 QY 439 AAAAGCCAAAACCTCTCTGCTTGTAGGCGCCATCGTCCAGGCTCTTCCATGAGTGCT 498  
 |||||  
 Db 409 AAGTCCAATAAACCAATCGTCTGATTGGCGCACAGCGCAACGCTCGGTTTCAGACTTT 350  
 |||||  
 QY 499 GATGGCCCCATGAATCTTATACGCGGTGAATGTAGCGATCAACAAGCTCTACTTAAC 558  
 |||||  
 Db 349 GACGGTCCCGCAACTTGTCTTAACGCTGTGGCATCGCTTGTATCCCAATCCCAAGGAT 290  
 |||||  
 QY 559 AAAGGAGTGGTGATTGTGATGAACGATGAGATTTCACGCCCGCAGAGAGCGCAAGCTC 618  
 |||||  
 Db 289 CGTGGCGCATCTGCTGGGATGAACAACAGATCAACGCTACCAAGTATGTACCACAGACA 230  
 |||||  
 QY 619 AACACACCGCAGTCAATGCAATTT-GCTTGGCCCAACACAGGTAAATTCGACAGTCTA 677  
 |||||  
 Db 229 CACACTGCCCAAGCTTGAACACTTTTAAAGTTCTGGTGAGTTCGGCTTTTATGGTGAGGTCTA 170  
 |||||  
 QY 678 TTATGGCAAGTCGAGTATTTCACTCAATCCGTTTCGACCTCACACCCCTTGAAGTGAGTT 737  
 |||||  
 Db 169 CCCAGACAGGTCGTTTACCACACTAGCGCATCGCAATTTAAACTTTCTCTATCAAGGC 110  
 |||||  
 QY 738 TGAT 741  
 |||||  
 Db 109 AGAT 106  
 |||||

RESULT 15  
 AQ873892/c  
 LOCUS  
 DEFINITION V83A5 mTn-3xHA/lacZ Insertion Library, strain AB972 Saccharomyces  
 cerevisiae genomic 5', genomic survey sequence.

ACCESSION  
 VERSION AQ873892.1 GI:6286136  
 KEYWORDS  
 SOURCE  
 ORGANISM

Saccharomyces cerevisiae (baker's yeast)  
 Saccharomyces cerevisiae  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE  
 AUTHORS

1 (bases 1 to 560)  
 Rosas-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,  
 desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniat, D., Jansen, R.,  
 Umansky, L., Heideman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,  
 Hager, K., Miller, P., Roeder, G.S. and Snyder, M.  
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and  
 Gene Disruption

JOURNAL  
 COMMENT

Contact: Kumar A  
 Michael Snyder, Dept. of Mol. Cell. and Dev. Biology  
 Yale University  
 P.O. Box 208103, New Haven, CT 06520-8103, USA  
 Tel: 203 432 9949  
 Fax: 203 432 6161  
 Email: anuj.kumar@yale.edu

te of mTn-3xHA/lacZ insertion.  
 Seq primer: GGCCTTCTTCTTTGGAAGTAC  
 Class: transposon-tagged.  
 Location/Qualifiers  
 1..560

FEATURES  
 source  
 /organism="Saccharomyces cerevisiae"  
 /mol\_type="genomic DNA"  
 /strain="AB972 - trp1 r(0) (S288C background)"  
 /db\_xref="taxon:4932"  
 /lab\_host="E. coli"  
 /clone\_lib="mTn-3xHA/lacZ Insertion Library, strain AB972"

/note="Vector: pHS6-Sal; A yeast genomic DNA library was prepared in pHS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtn-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT		155 a	127 c	119 g	158 t	1 others
ORIGIN						
Query Match						
Best Local Similarity 7.7%; Score 87; DB 28; Length 560;						
Matches 177; Conservative 54.0%; Pred. No. 1.8e-14;						
Matches 177; Conservative 0; Mismatches 151; Indels 0; Gaps 0;						
Qy	383	TCACCCATGGA	CTGACACCATG	GAAGAGACCG	CTTCTTCCTCA	ACCTCACGGTGA
Db	503	TCACCTCATGG	ACCGGACACTAT	GGAGAGACGCT	ATTTCTTAGAT	TTTGACCATANAT
Qy	443	GCCAAAAAC	CTGCTGCCCTT	GTAGGCGCCAT	GGTCCAGGCTC	TTCCATGAGTGC
Db	443	CAGAGAGAG	CCAGTATGAT	TCGAGGGGAT	ATGGTCCAGCA	CTGCCACGTCTG
Qy	503	GCCCCATGA	ATCTCTATA	AGCCGCGTGA	ATGAGCGATCA	ACAAAGCCTCTA
Db	383	GCCCAATGA	ATTTATATCA	AGCAGTGTCT	ATTGCTTCTG	AGAAATCACTGG
Qy	563	GAGTGTGAT	TGTGATGAG	ATGAGATTC	ACGCCGCCAG	AGAGCGACCAAG
Db	323	GCACGATGA	TCACTCTAA	ACGATCGTAT	TGCCCTCTGG	GTGTTGGACAA
Qy	623	CCACCGCAG	TCAATGCTT	TCGCCCAACA	CACAGGTAA	AAATCGGCACAG
Db	263	CCNACTCT	TTTAGATACA	TTCAGAGCGG	ATGATTTAG	GTACTTTTCAAT
Qy	683	GCAAGTCG	AGATTTTCA	CTCAATCCGT	710	
Db	203	ATGACGTG	GAGTTTACT	TACCCACCAG	T 176	

Search completed: August 27, 2003, 09:30:48  
Job time : 3478.96 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:28:33 ; Search time 127.739 Seconds  
(without alignment)  
9607.814 Million cell updates/sec

Title: US-09-937-982-1  
Perfect score: 30  
Sequence: 1 tccggatccagcgcctctgtttgatggct 30

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2898711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: gb.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	30	100.0	30	6	AR157942	AR157942 Sequence
2	30	100.0	30	6	AX464425	AX464425 Sequence
3	28	93.3	1133	6	AR157944	AR157944 Sequence
4	28	93.3	1133	6	AX464427	AX464427 Sequence
5	21.6	72.0	119959	2	AP005865	AP005865 Oryza sat
6	21.6	72.0	125505	2	AP004177	AP004177 Oryza sat
7	21.6	72.0	316050	1	AX321859	AX321859 Nitrosomo
8	21.2	70.7	146405	2	AC143437	AC143437 Macaca mu
9	21.2	70.7	221887	9	AL158052	AL158052 Human DNA
10	21	70.0	1035	1	WSANSAGEN	WSANSAGEN
11	21	70.0	2505	1	WSDCUANSA	WSDCUANSA
12	21	70.0	3606	1	WSAJ2933	WSAJ2933 Wolinella
13	20.8	69.3	4927	1	AE010450	AE010450 Methanopy
14	20.6	68.7	908	9	HS340724	HS340724 Homo sapi
15	20.6	68.7	13438	1	AE014767	AE014767 Bifidobac
16	20.6	68.7	172378	2	AC134210	AC134210 Rattus no
17	20.6	68.7	249207	2	AC102734	AC102734 Mus muscu
18	20.6	68.7	299950	1	AP005372	AP005372 Thermosyn
19	20.6	68.7	349980	6	AX492786	AX492786 Sequence
20	20.6	68.7	349980	6	AX553953	AX553953 Sequence
21	20.4	68.0	781	1	ECECODXXI	ECECODXXI E.coli DNA
22	20.4	68.0	781	1	ECECOPRI	ECECOPRI E.coli DNA
23	20.4	68.0	2200	1	ECDNAHSDM	ECDNAHSDM
24	20.4	68.0	4343	7	PPIPHDDOC	PPIPHDDOC
25	20.4	68.0	16228	8	F6113	F6113
26	20.4	68.0	50015	8	AB025611	AB025611 Arabidops
27	20.4	68.0	89163	9	HS1009H6	HS1009H6 Human DNA
28	20.4	68.0	137938	2	AL928940	AL928940 Mus muscu
29	20.4	68.0	207823	10	AL929142	AL929142 Mouse DNA
30	20.2	67.3	173354	9	HS149A16	HS149A16 Human DNA
31	20	66.7	157493	9	AC069248	AC069248 Homo sapi
32	20	66.7	160667	9	AC024157	AC024157 Homo sapi
33	20	66.7	180413	2	AC119169	AC119169 Mus muscu
34	20	66.7	254721	2	AC120733	AC120733 Rattus no
35	20	66.7	255406	2	AC133099	AC133099 Mus muscu
36	20	66.7	268801	2	AC126068	AC126068 Rattus no
37	20	66.7	305735	2	AC134916	AC134916 Mus muscu
38	20	66.7	343010	2	AC120175	AC120175 Mus muscu
39	19.8	66.0	6642	6	I64799	I64799 Sequence 5
40	19.8	66.0	6840	6	AX377877	AX377877 Sequence
41	19.8	66.0	6840	9	HS1R11	HS1R11 H sapiens m
42	19.8	66.0	6843	6	E13292	E13292 cDNA encodi
43	19.8	66.0	6843	6	I64800	I64800 Sequence 6
44	19.8	66.0	6902	6	AX098219	AX098219 Sequence
45	19.8	66.0	6981	6	AX336385	AX336385 Sequence

# ALIGNMENTS

RESULT 1	AR157942	Sequence 1 from patent US 6251388.	30 bp	DNA	linear	PAT 17-OCT-2001
LOCUS	AR157942	Sequence 1 from patent US 6251388.	30 bp	DNA	linear	PAT 17-OCT-2001
DEFINITION	AR157942	Sequence 1 from patent US 6251388.	30 bp	DNA	linear	PAT 17-OCT-2001
ACCESSION	AR157942	Sequence 1 from patent US 6251388.	30 bp	DNA	linear	PAT 17-OCT-2001
VERSION	AR157942.1	GI:16219884	30 bp	DNA	linear	PAT 17-OCT-2001
KEYWORDS	Unknown.	Unknown.	30 bp	DNA	linear	PAT 17-OCT-2001
SOURCE	Unknown.	Unknown.	30 bp	DNA	linear	PAT 17-OCT-2001
ORGANISM	Unknown.	Unknown.	30 bp	DNA	linear	PAT 17-OCT-2001
REFERENCE	1 (bases 1 to 30)	Unclassified.	30 bp	DNA	linear	PAT 17-OCT-2001
AUTHORS	Durden,D.L.	Durden,D.L.	30 bp	DNA	linear	PAT 17-OCT-2001
TITLE	Utilization of Wolinella succinogenes asparaginase to treat diseases associated with asparagine dependence	Utilization of Wolinella succinogenes asparaginase to treat diseases associated with asparagine dependence	30 bp	DNA	linear	PAT 17-OCT-2001
JOURNAL	Patent: US 6251388-A 1 26-JUN-2001;	Patent: US 6251388-A 1 26-JUN-2001;	30 bp	DNA	linear	PAT 17-OCT-2001

\* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced	
* by the finished sequence as soon as it is available and	
* the accession number will be preserved.	
FEATURES	Location/Qualifiers
	1..119959
/organism="Oryza sativa (japonica cultivar-group)"	
/mol_type="genomic DNA"	
/cultivar="Nipponbare"	
/db_xref="taxon:39947"	
/chromosome="2"	
/clone="OSUNBB0037J12"	
BASE COUNT	34012 a 26175 c 25976 g 33746 t 50 others
ORIGIN	
Query Match	
Best Local Similarity 72.0%; Score 21.6; DB 2; Length 119959;	
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	3 CGGATCCAGCGCTCTGTTTGATGGCT 30
DB	109538 CGCCTCTCTCTCTCTGTTTGATGGCT 109565
RESULT 6	
LOCUS	AP004177/c 125505 bp DNA linear HTG 21-MAR-2002
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 2 clone
ACCESSION	OJ1116 E03, *** SEQUENCING IN PROGRESS ***.
VERSION	AP004177.1 GI:15718434
KEYWORDS	HTG; HTGS PHASE2.
SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
Ehrhartoideae; Oryzeae; Oryza.	
REFERENCE	
1 Sasaki, T., Matsumoto, T. and Yamamoto, K.	
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC	
clone:OJ1116 E03	
Published Only in Database (2001)	
REFERENCE	
2 (bases 1 to 125505)	
Sasaki, T., Matsumoto, T. and Yamamoto, K.	
Direct Submission	
Submitted (20-SEP-2001) Takuji Sasaki, National Institute of	
Agrobiological Sciences, Rice Genome Research Program; Kamondai	
2-1-2, Tsukuba, Ibaraki 305-8602, Japan	
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,	
Tel:81-298-38-7441, Fax:81-298-38-7468)	
The nucleotide sequence of this BAC clone was generated by	
combining Monsanto and RGP-Japan sequencing data.	
NOTE: It currently consists of 1 contigs. Gaps between the contigs	
are represented as runs of N. The order of the pieces is believed	
to be correct as given, however the sizes of the gaps between them	
are based on estimates that have provided by the submitter. This	
sequence will be replaced by the finished sequence as soon as it is	
available and the accession number will be preserved.	
* NOTE: this is a 'working draft' sequence.	
* This sequence will be replaced	
* by the finished sequence as soon as it is available and	
* the accession number will be preserved.	
FEATURES	Location/Qualifiers
	1..125505
/organism="Oryza sativa (japonica cultivar-group)"	
/mol_type="genomic DNA"	
/cultivar="Nipponbare"	
/db_xref="taxon:39947"	
/chromosome="2"	
/clone="OJ1116 E03"	
BASE COUNT	33795 a 27965 c 29679 g 34012 t 54 others
ORIGIN	
Query Match	
72.0%; Score 21.6; DB 2; Length 125505;	

Best Local Similarity 85.7%; Pred. No. 57;	
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	3 CGGATCCAGCGCTCTGTTTGATGGCT 30
DB	79713 CGCCTCTCTCTCTCTGTTTGATGGCT 79686
RESULT 7	
LOCUS	EX321859 316050 bp DNA linear BCT 23-APR-2003
DEFINITION	Nitrosomonas europaea ATCC 19718, complete genome; segment 4/10.
ACCESSION	EX321859 AL954747
VERSION	EX321859.1 GI:30138466
KEYWORDS	complete genome.
SOURCE	Nitrosomonas europaea ATCC 19718
ORGANISM	Nitrosomonas europaea ATCC 19718
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;	
Nitrosomonadaceae; Nitrosomonas.	
REFERENCE	
1 Chain, P., Lamerdin, J., Larimer, F., Regala, W., Land, M., Hauser, L.,	
Hooper, A., Klotz, M., Norton, J., Sayavedra-Soto, L., Arciero, D.,	
Homes, N., Whittaker, M. and Arp, D.	
Complete Genome Sequence of the Ammonia-Oxidizing Bacterium and	
Obligate Chemolithoautotroph Nitrosomonas europaea	
J. Bacteriol. 185 (9), 2759-2773 (2003)	
12700255	
REFERENCE	
2 (bases 1 to 316050)	
Larimer, F.	
Direct Submission	
Submitted (12-NOV-2002) Submitted on behalf of the Nitrosomonas	
genome consortium, the DOE Joint Genome Institute, Production	
Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598,	
USA, and the Genome Analysis Group, Oak Ridge National Laboratory,	
1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;	
larimerf@ornl.gov	
Nitrosomonas genome consortium	
FEATURES	
Location/Qualifiers	
1..316050	
/organism="Nitrosomonas europaea ATCC 19718"	
/mol_type="genomic DNA"	
/strain="ATCC 19718"	
/db_xref="taxon:228410"	
complement (639..1223)	
/locus_tag="NE0846"	
complement (639..1223)	
/locus_tag="NE0846"	
/codon_start=1	
/evidence=not experimental	
/transl_table=11	
/product="conserved hypothetical protein"	
/protein_id="CAD84757.1"	
/db_xref="GI:30138467"	
/translation="MIRNNSFYRIDPLKSPIKANPLRNPGESLNHLQNFIDGIV	
PYFTAAFCVLIAAEWIRWYOTSPNPVFTVMAIPCIATLWKIYKGRKEVKRIKL	
GIAGELAVQFLERLRAGQSHIFHDIPKGFNLHVHSSGIIVETKLSKPDGRGE	
SKLYVDGNHLKNSLTALDNFITTQVANSRWLRE"	
order (complement (1038..1094), complement (951..1007))	
/locus_tag="NE0846"	
/note="2, probable transmembrane helices predicted by	
TMHMM2.0"	
complement (1351..2301)	
/locus_tag="NE0847"	
complement (1351..2301)	
/locus_tag="NE0847"	
/function="COGS COG0726"	
/codon_start=1	
/evidence=not experimental	
/transl_table=11	
/product="conserved hypothetical protein"	
/protein_id="CAD84758.1"	
/db_xref="GI:30138468"	
/translation="WKILPSLPSPAGKNSKLSIVYHRLPEPDLITGEGGIAQFEKG	

LPASGVAKLWLPDASDTVYQTQGNNSGKATSAFYMLPCTQSQVFPALWQKGE  
 ERTVRSVVKTAORTVDHLYSPAKISIPONTQYRLOSTKTHIPLDVWVKMARSVTV  
 SPANAHTVLAQRAAIYDWVNNVAYDINGRQSGDRLSLSRKENISGKCGDHALF  
 VGLARASGIPARIQVGRISDSALNKNLKGSGDISTSQHQAEFYLSGLGWVPVNPSP  
 VARVMAASLSGHNHSDISOLREKLFSGWEMNWAFNERENVDLGTTKVSAAISKLPFF  
 GYPYAEIDGRLOQDSLPESFSYKITSVLVGTGAKL  
 complement (5615..7336)  
 /gene="yv9Q"  
 /locus\_tag="NE0852"  
 complement (5615..7336)  
 /gene="yv9Q"  
 /locus\_tag="NE0852"  
 /EC\_number="1.8.1.2"  
 /function="InterPro IPR000660:IPR005117 COGs COG0155"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /transl\_table=11  
 /product="Nitrite and sulfite reductase"  
 /protein\_id="CAD84763.1"  
 /db\_xref="GI:30138473"  
 /translation="MANTLPETDRSDISQPLRLSPDSLKAESDYLRGTTIALGLLDD  
 RITSAPGNDIKLWKHGIYEQDREIRDERRRRQKLEPAFQFMIRVLPGICCTTERW  
 LKISELACAGNETLMTATKQTQTFHWWLQKNIVPLIRGHEVLLDVTAAAGDSDRGV  
 MATVNPQPALQALAAKATLSDSHVIPKTRAYHEIYGEERIASSEPEEPFYGTQYM  
 PRKFGIPVPPNNIDITYAQDILGYTAIIENGKIAGFNVAIGGGMGRTDKAPHTYPR  
 TASVIGITFDRLNIDVSTEAVGQVQDRGNADRSRARKFYTTDDKGLDWIKLAIEDRA  
 GPLSEARPYDFTNADYIWEISGDGFHFHTLFIENGRNLNRLMSKALAOIAHVHKGHF  
 RLTPNQMLYANATADKPIEALLRETGLIAFNERSVLRLNSMCVAPLTCGLAMAD  
 SERVPLDITVKIGILTRYNLONEPITLRMTGCPNGCSRPFYAEIGLTRAPGKYNLY  
 LGGFGFQRLNRLYRNIENGPALLEINELVGLRYATERLPDSEHFGDFTIRAGIIREVT  
 EGRFSND"  
 complement (7345..9180)  
 /gene="yv9R"  
 /locus\_tag="NE0853"  
 complement (7345..9180)  
 /gene="yv9R"  
 /locus\_tag="NE0853"  
 /EC\_number="1.8.1.2"  
 /function="InterPro  
 IPR001094:IPR001226:IPR001433:IPR001709:IPR003097 COGs  
 COG0369"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /transl\_table=11  
 /product="Sulfite reductase flavoprotein subunit"  
 /protein\_id="CAD84764.1"  
 /db\_xref="GI:30138474"  
 /translation="MNASQLSRGTLTAEQWLVEALGSSLTPEHAYWISGYFAGLGS  
 LRRSSATQSDPVCIPPVYQAEPLVAARSLTILHISETGNSTELAIKLAALAVEOGL  
 SPTLVGIADYVKRLKEODLLIITSTHEGDPQSGMEFFFEVGRKAPSLGRLVA  
 ILALGMSYHDFRCAGKRLDERFEALGAKRLQPPVCDVDYEDPAAVSTGILALLAA  
 EQASATSSVAPSKTGTGTOQANNSVYKRNPPFATVIDNIVLSRGSGSTRTHHIEIS  
 LADSGITPEFGALGIAHPANDPAVEALLANMNPDPATIVKGGTITFLFDALGKSF  
 ITATRLFDHGEUETGTTDIPSAMNDTKRAAFLDHNDHIIDIIRAYPLKGLDPPQRFVD  
 GURLOPRLYSIASLSICPDASHITIAIPVRYKLHGEPRSGVAGSHADRAIDSVLP  
 VYIQSNPRLPSNGSEPTIMIGATGVAPYAFRAFMQERANGSGWMPFGRNFRFTDF  
 LYQTEQWDLKNGVLTKMEVASRDCAEKVVQHRIYEHADVVAVLVEEGAYIYLGD  
 ASHMAPDVHNLIVTVIEBQASVOREAEAYRLDLDQDRGYQRDYY"  
 complement (9708..10664)  
 /gene="cysB"  
 /locus\_tag="NE0854"  
 complement (9708..10664)  
 /gene="cysB"  
 /locus\_tag="NE0854"

```

RESULT 8
AC143437
LOCUS
DEFINITION
AC143437
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Macaca mulatta clone CH250-267L24, *** SEQUENCING IN PROGRESS ***.
HTG; HTGS PHASE2; HTGS_PGI.
Macaca mulatta (rhesus monkey)
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
1 (bases 1 to 146405)
Csuros.M. and Milosavljevic.A.
Pooled genomic indexing (PGI): mathematical analysis and experiment
design
(in) Guigo,R. and Gusfield,D. (Eds.);
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
2 (bases 1 to 146405)
Milosavljevic.A., Sodergren,E., Csuros,M., Li,B., Jackson,A.R.,
Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L.,
Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbakid,J.,
Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Egan,A., Earhart,C., Edwards,C.C.,
Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsif,F.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondei,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Ogih,M., Okuwon,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojebokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shoostari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A.,
Wleznyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
3 (bases 1 to 146405)
Worley,K.C.
Direct Submission
Submitted (09-APR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center

```

```

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: LBPH
Center clone name: CH250-267L24
----- Summary Statistics
Chemistry: Dye-primer Bodipy: inf% of reads
Chemistry: Dye-terminator Big Dye: inf% of reads
Consensus quality: 4889 bases at least Q40
Consensus quality: 5926 bases at least Q30
Consensus quality: 6870 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: The contigs are based on the application
* of the PGI method using the Human genome (NCBI build 31)
* as the comparative genome.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 146405: contig of 146405 bp in length.
* Location/Qualifiers
* 1..146405
* /organism="Macaca mulatta"
* /mol_type="genomic DNA"
* /db_xref="taxon:9544"
* /clone="CH250-267L24"
* 1..146405
* /note="assembly_name:CH250-267L24.1B
* CONFIDENCE:0.67"
BASE COUNT 2409 a 2871 c 2652 g 2518 t 135955 others
ORIGIN
Query Match 70.7%; Score 21.2; DB 2; Length 146405;
Best Local Similarity 88.5%; Pred. No. 88;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCCGGATCCAGCGCTCTGTTTGTAT 26
Db 95612 TCCGGATCCAGCGCTCTGTTTGTAT 95637
-----
RESULT 9
AL158052
LOCUS
DEFINITION
AL158052
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AL158052
Human DNA sequence from clone RP11-570J20 on chromosome
9q33.1-34.12, complete sequence.
AL158052
AL158052.10 GI:12964258
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 221887)
Direct Submission
Submitted (05-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Feb 19, 2001 this sequence version replaced gi:12657095.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission

```

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr9>  
RP11-570J20 is from the library RPC1-11.2 constructed by the group of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-570J20 The true left end of clone RP11-277p3 is at 128358 in this sequence.

#### FEATURES

Location/Qualifiers  
1..221887  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="9"  
/map="q33.1-34.12"  
/clone="RP11-570J20"  
/clone\_lib="RPC1-11.2"  
18060..18125

#### misc\_feature

/note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."

#### misc\_feature

/note="Tandem repeat. Forced join. Gap size estimated to be approximately 225bp by restriction digest data."

#### BASE COUNT

ORIGIN

Query Match 70.7%; Score 21.2; DB 9; Length 221887;  
Best Local Similarity 88.5%; Pred. No. 87;  
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TCCGATCCAGCGCTCTGTTTGCAT 26

Db 107447 TCCGGAACCTGCGCTCTGTTTGCAT 107472

#### RESULT 10

WSANSAGEN 1035 bp DNA linear BCT 01-APR-1995  
LOCUS W.succinogenes ansa gene.  
DEFINITION X83689  
ACCESSION X83689  
VERSION X83689.1 GI:758651  
KEYWORDS ansa gene; asparaginase.  
SOURCE Wolinella succinogenes  
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Wolinella.

#### REFERENCE

1 Derst, C. and Roehm, K.H.  
AUTHORS Cloning and sequencing of L-asparaginase from Wolinella succinogenes  
TITLE Unpublished  
JOURNAL  
REFERENCE 2 (bases 1 to 1035)  
AUTHORS Roehm, K.H.  
TITLE Direct Submission  
JOURNAL Submitted (05-JAN-1995) K.H. Roehm, Physiologische Chemie der Philipps-Univ., Karl-von-Frisch-Str.1, D-35033 Marburg (Lahn), FRG

#### FEATURES

source

#### Location/Qualifiers

1..1035  
/organism="Wolinella succinogenes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:844"  
/clone\_lib="lambda EMBL3"  
20..1012  
/gene="ansa"  
20..1012  
/gene="ansa"  
/EC number="3.5.1.1"  
/codon\_start=1  
/transl\_table=11  
/product="L-asparaginase"  
/protein\_id="CAA58658.1"  
/db\_xref="GI:758652"  
/translation="MAKPOVTILATGCTIAGSGSSVKSSYSAGAVTVDKLAAPALNDLTIKGEQISSIGSOEMTKVWLKLRVNEILAQKETEAVIITHGDTMETTAFPLNLTIVSKQKPVVLGAMRPGSSMSADGPMNLVNAVINKASTNGKVVIVMDEIHAAREATKLTAVNAFASPTKIGTIVYKVEYFTQVRPHLTASEFIDISKIEIPRV DILYAHDDTDVLNAAALQAGAKGIHAGMGNGNPFPLTQNALEKAAGSGVVARSSRVSGSSTTQEAEDDDKKLGFVATESLNPKARVLLMLALTKTSDREAQIKIFSTY"

#### BASE COUNT

ORIGIN

Query Match 70.0%; Score 21; DB 1; Length 1035;  
Best Local Similarity 100.0%; Pred. No. 1;le+02;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGCGCTCTCTGTTTGTGCGCT 30

Db 5 AGCGCTCTCTGTTGCGCT 25

#### RESULT 11

WSDCUANSA 2505 bp DNA linear BCT 18-JUL-1997  
LOCUS W.succinogenes dcua and ansa genes.  
DEFINITION X89215  
ACCESSION X89215  
VERSION X89215.1 GI:895917  
KEYWORDS ansa gene; asparaginase; C4-dicarboxylate membrane transporter; dcua gene.  
SOURCE Wolinella succinogenes  
ORGANISM Wolinella succinogenes  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Wolinella.

#### REFERENCE

1 Lubkowski, J., Palm, G.J., Gilliland, G.L., Derst, C., Rohm, K.H. and Wlodawer, A.  
AUTHORS Crystal structure and amino acid sequence of Wolinella succinogenes L-asparaginase  
TITLE  
JOURNAL Eur. J. Biochem. 241 (1), 201-207 (1996)  
MEDLINE 97054610  
PUBMED 8898907  
REFERENCE 2 (bases 1 to 2505)  
AUTHORS Roehm, K.H.  
TITLE Direct Submission  
JOURNAL Submitted (27-JUN-1995) K.H. Roehm, Institut fuer Physiologische Chemie, Philipps-Universitaet Marburg, Karl-von-Frisch-Str.1, 35033 Marburg, FRG

#### COMMENT

Related sequences X83689.

#### FEATURES

source

#### Location/Qualifiers

1..2505  
/organism="Wolinella succinogenes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:844"  
/clone\_lib="lambda EMBL3"  
1..1068  
/gene="dcua"  
1..1068  
/gene="dcua"  
/codon\_start=1

#### gene

#### CDS

```

/transl_table=11
/product="C4-dicarboxylate membrane transporter"
/protein_id="CAA61502.1"
/db_xref="GI:895918"
/db_xref="SWISS-PROT:O34245"
/translation="KLESPKQKHTFLAPVVTYFMTLLAGTGTAFSTLPIAKVAKE
QGVKSVLEVCGRGLSIAITASPVSAVPMGRVDELGVGLQLLAICTPTTFCG
MTATICNLNFNDLSDQPIQERVAKGKRLGQTOYVAKGAKLSVILFILGLAVV
FYATISKNVGLIOPVGRDSAINMFMLTATLIAMFAKIDTDAVLKASTFKSGMTA
CTICVLGAWLGDFTVSNHNDIKALSGGILEGHPWMLAILLFFASMLLYSOATAKAL
IPAAPLSDVPVSLIASAASGLFVLPTPTLLAAVQMDTGTSTRICKFVFNHPFV
PGVLAIGISVALGFIIVAPILL"
misc_feature
1069..1188
/feature="intergenic region"
terminator
1079..1102
/feature="potential rho-independent terminator for dcua"
RBS
1171..1177
/feature="putative"
gene
1189..2181
/feature="ansa"
CDS
1189..2181
/feature="ansa"
/EC_number="3.5.1.1"
/codon_start=1
/transl_table=11
/product="asparaginase"
/protein_id="CAA61503.1"
/db_xref="GI:895919"
/db_xref="SWISS-PROT:P50286"
/translation="MAKPOVITLGTGTAGSGSSVSKSSAGAVTVDKLLAAVPAI
NDLTKGQISIGSQMGRTGKVLKRVNLLAOKETRAVITHTGDTMEETAFF
LNLTKSQRPVILVGNMRGSSMSADGPMNLVAVNVAINKSTNGVVIWNDEIHA
AREATKNTAVNAFSPNTGKIGTYGKVEYFTQSVRPHTLASEPDISKIELPRV
DILYAPQDTEVLVNAALQAGAKIIGHAGNGNPPPLTONALEKAAKGGVVAKRSR
VSGSSTQAEVDDKKLGFVATESLNPQKARVLLMLTKTSDREAIQKIPSTY"
2186..2219
terminator
/feature="potential rho-independent terminator for ansa"
BASE COUNT 594 a 690 c 553 g 668 t
ORIGIN
70.0%; Score 21; DB 1; Length 2505;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGCGCCTCTGTTTGTGGCT 30
|||||
Db 1174 AGCGCCTCTGTTTGTGGCT 1194

RESULT 12
WSAJ2933 3606 bp DNA linear BCT 05-OCT-2000
LOCUS Wolinella succinogenes aspa, dcua genes and partial ansa gene.
DEFINITION
AJ002933
ACCESSION
VERSION AJ002933.1 GI:2644958
KEYWORDS
ansa gene; aspa gene; aspartate ammonia-lyase; C4-dicarboxylate
membrane transporter; dcua gene; L-asparaginase.
SOURCE
Wolinella succinogenes
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Wolinella.

REFERENCE
1 Ullmann, R., Gross, R., Simon, J., Unden, G. and Kroger, A.
Transport of C(4)-dicarboxylates in Wolinella succinogenes
J. Bacteriol. 182 (20), 5757-5764 (2000)
20461222
MEDLINE
PUBMED 11004174
REFERENCE
2 (bases 1 to 3606)
Ullmann, R.
Direct Submission
Submitted (19-NOV-1997) Ullmann R., Institut fuer Mikrobiologie,
Johann-Wolfgang-Goethe Universitaet, Biozentrum N240,
Marie-Curie-Strasse 9, 60439 Frankfurt, GERMANY

```

```

COMMENT Related sequence: X89215.
FEATURES
source Location/Qualifiers
1..3606
/organism="Wolinella succinogenes"
/mol_type="genomic DNA"
/db_xref="taxon:844"
400..1815
/feature="aspA"
gene
400..1815
/feature="aspA"
CDS
/codon_start=1
/transl_table=11
/product="aspartate ammonia-lyase"
/protein_id="CAA05764.1"
/db_xref="GI:2644959"
/db_xref="SPTREMBL:O34244"
/translation="MASTRIEHDLIGDLFVNDYCYGTQARAKENFHLISGVLLSSMP
TFIASLAKVKKAAALANFELGLLDEKIKNAICEACDLIIAGKHDFQVVDALQGGAGT
STNNANEVIANVALELMGHKKGEYKHPNNHNLQSNDNAYPTAFRVALYEKLEIE
LTVSMGILRDSFAKAEFKDIIKMGRTQLQDAVPMTLGQEFKYAVMMTIDTDRVLE
ARNLVREMLGATAIGTINSHPDYATLVKKLQEVTVGRPFITASDLVEATDITGAYV
QISGLKRVSTKVSUKIANDLRLSSGRAGLSEINLPPMQGSSIMPGKVNPIPEVV
NQCYHVGADMTVTIASGGQLQLVNPFVPVIAFSLFSSINMMRAFEFLAEKCVNGI
TANADNCRDYLKISIGLITNALNPYIGYENSTSVAKEALESCKSVHDIIVLERKLLSKEE
LEEIIRPNMIKPKVVTLTKK"
1828..3129
/feature="dcua"
gene
1828..3129
CDS
/feature="dcua"
/codon_start=1
/transl_table=11
/product="C4-dicarboxylate membrane transporter"
/protein_id="CAA05765.1"
/db_xref="GI:2644960"
/db_xref="SWISS-PROT:O34245"
/translation="MLVFEILVVAIAFLVRLGGIAGYAGGLGMIVLCGLGLKPG
SIFDVIILIMSVAIAAMQVAGGLDYLHLAEKLRISQPKHITFLAPVVTFTMILL
AGTGTAFSTLPVIAKVEQGVKPSVLSIAVASQIAITASPSAAVFMGSALEP
LGVYQLLALICITPTFIGMTAFICNLTDLSDKDPYQERVAKGMLRGQTQV
AKPKALSVILFLVGLAVFYATVATSKNGLIPIVGRDSAIMFMMLTATLIAMF
AKITDAVLNASTFKSGMTACICVLGVALGDTFVSNHNDIKGFAGGILEGHPWMLA
ITLFFASMLLYSQARTAKALIIPAAIALSDVPVLLIASFPAVSALFVLPTPTLLAAVQ
MDTGSTRIKGFVFNHPFIVPGVLAIGISVALGFIIVAPILL"
3250..3606
/feature="ansa"
gene
3250..3606
CDS
/feature="ansa"
/codon_start=1
/transl_table=11
/product="L-asparaginase"
/protein_id="CAA05766.1"
/db_xref="GI:2644961"
/translation="MAKPOVITLGTGTAGSGSSVSKSSAGAVTVDKLLAAVPAI
NDLTKGQISIGSQMGRTGKVLKRVNLLAOKETRAVITHTGDTMEETAFF
LNLTKSQRPVILVGNMRGSSMSADGPMNLVAVNVAINKSTNGVVIWNDEIHA
AREATKNTAVNAFSPNTGKIGTYGKVEYFTQSVRPHTLASEPDISKIELPRV
DILYAPQDTEVLVNAALQAGAKIIGHAGNGNPPPLTONALEKAAKGGVVAKRSR
VSGSSTQAEVDDKKLGFVATESLNPQKARVLLMLTKTSDREAIQKIPSTY"
2186..2219
terminator
/feature="potential rho-independent terminator for ansa"
BASE COUNT 919 a 973 c 806 g 908 t
ORIGIN
70.0%; Score 21; DB 1; Length 3606;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGCGCCTCTGTTTGTGGCT 30
|||||
Db 3235 AGCGCCTCTGTTTGTGGCT 3255

RESULT 13
AE010450
LOCUS Methanopyrus kandleri AV19 section 149 of 157 of the complete
DEFINITION genome.
ACCESSION AE010450 AE009439

```



SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 908)  
 AUTHORS Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, P., Kvasha, S.M.,  
 Podowski, R.M., Matushkin, Y.G., Ganchandani, A., Muravenko, O.V.,  
 Levitsky, V.G., Kolchanov, N.A., Protodopov, A.I., Kashuba, V.I.,  
 Kisselev, L.L., Wasserman, W., Wahlestedt, C., and Zabarovsky, E.R.  
 TITLE NotI flanking sequences: a tool for gene discovery and verification  
 of the human genome  
 JOURNAL Nucleic Acids Res. 30 (14), 3163-3170 (2002)  
 MEDLINE 22131767  
 PUBMED 12136098  
 REFERENCE 2 (bases 1 to 908)  
 AUTHORS Zabarovsky, E.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-MAY-2001) Microbiology and Tumourbiology Centre,  
 Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,  
 Sweden  
 FEATURES  
 source Location/Qualifiers  
 1. 908  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="NL1-ZC2RS"  
 BASE COUNT 212 a 272 c 252 g 169 t 3 others  
 ORIGIN  
 Query Match 68.7%; Score 20.6; DB 9; Length 908;  
 Best Local Similarity 85.2%; Pred. No. 1.7e+02;  
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 CCGGATCCAGCGCCCTCTTTGATGG 28  
 |||||  
 Db 353 CCGGAGCCAGCGCCCTCTGATGGATGG 379  
 RESULT 15  
 AE014767  
 LOCUS  
 DEFINITION Bifidobacterium longum NCC2705 section 154 of 202 of the complete  
 genome.  
 ACCESSION AE014767  
 VERSION AE014295  
 KEYWORDS  
 SOURCE  
 ORGANISM Bifidobacterium longum NCC2705  
 Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;  
 Bifidobacteriaceae; Bifidobacterium.  
 REFERENCE 1 (bases 1 to 13438)  
 AUTHORS Schell, M.A., Karmirantzou, M., Snel, B., Vilanova, D., Berger, B.,  
 Pessi, G., Zwielen, M.-C., Desiere, F., Bork, P., Delley, M.,  
 Pridmore, D., and Arigoni, F.  
 TITLE The genome sequence of Bifidobacterium longum reflects its  
 adaptation to the human gastrointestinal tract  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14422-14427 (2002)  
 PUBMED 12381787  
 REFERENCE 2 (bases 1 to 13438)  
 AUTHORS Schell, M.A., Karmirantzou, M., Snel, B., Vilanova, D., Berger, B.,  
 Pessi, G., Zwielen, M.-C., Desiere, F., Bork, P., Delley, M.,  
 Pridmore, D., and Arigoni, F.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-AUG-2002) Bioscience, Nestle Research Center, P.O.  
 Box 44, Lausanne 26 1000, Switzerland  
 FEATURES  
 source Location/Qualifiers  
 1. 13438  
 /organism="Bifidobacterium longum NCC2705"  
 /mol\_type="genomic DNA"  
 /strain="NCC2705"  
 /db\_xref="taxon:206672"  
 95. .2941  
 /gene="BL1369"  
 /note="COG family: predicted transcriptional regulator  
 containing the HTH domain"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="AA25173.1"  
 /db\_xref="GI:23326644"  
 95. .2941  
 /gene="BL1369"  
 /note="COG family: ABC-type transport systems, involved in  
 lipoprotein release, ATPase components; PFAM\_ID: ABC\_tran;  
 PFAM ID: DUF214"  
 /codon\_start=1  
 /transl\_table=11  
 /product="large protein with N-terminal similarity to ATP  
 binding protein of ABC transporter"  
 /protein\_id="AA25170.1"  
 /db\_xref="GI:23326641"  
 /translation="MLQIKSISKRYKTCDFVQALDKVSLNRDSEFVALIPGSGSGK  
 TLLNIIGLDYDGDIVINGISTRQYKDRDSDSYRHNHTIGFVFQSVNLPHOTILS  
 NVELALITSGIKADRRERARKALEKGLGHEHINKKPNQSGQQRVAIRNLVNDP  
 DIVLADEPTGALDSQTSQIWMLLKDVAKDLVVMVTHNPELAQYATRIYLRDGI  
 RSDPFFVSTAEQAPVYKTMGRASMSFATSLALSNKLTAKRTLLTTFAGSIGI  
 IGIALISVSTGVNTYISDIORDTWTAPITVDSTQFLSSMMGQMGDDGYGKTKH  
 KTDGIYDPRSVKQASLITSTENNLTRFKKYLDSKSEIHOYVSGTIGIOYTVDVKE  
 SVFDHDPGTILVNDGVTIGSSDSASMASQVASTSSGMSGTSSITSOQMSMLTGKTL  
 ENAAPDSFNEIMPAGDSDSLGVKITDNYQVSWPKSKDEVLVLDDNNNSVPLTTL  
 YELGLLPASDIHEMMKSLNAGDKVSTPQDKIDYAKALDQTLTLMIPASQYVKGDDGHI  
 RFIGNDKDBIEORLETKLVGVVVKAKDASATPLAAGVGYARALNTDLIDRASS  
 ALVTDKANKPNTVINGMTFSPDDATKADATVYASLGVTQKANKMSWMSAQCS  
 GTGASDAAGAGVAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  
 YVSTGYDNLAAFGVVRSDRAPSSINIYADSFEDKDHADATDYNNTVSKDKITYT  
 ASKHNVSQVFNAETGIIIGLGLGLVGLTLLNIPINAVLHHFIGNADVNAALPVTGG  
 VVLVLSVVLTLIGGLIPSRKAQKQDPATLRTE"  
 3143. -4042  
 /gene="BL1370"  
 3143. -4042  
 /gene="BL1370"  
 /note="COG family: ATPases involved in chromosome  
 partitioning; PFAM\_ID: fer4\_NifH; PFAM\_ID: ParaA"  
 /codon\_start=1  
 /transl\_table=11  
 /product="widely conserved hypothetical protein in Para  
 family"  
 /protein\_id="AA25171.1"  
 /db\_xref="GI:23326642"  
 /translation="MTGTVVAPPANPDVDKVDSDMTDLGREGYETFPAPFLOQHG  
 AVIACNKGKGGVKTSSINAGLAQYGRVLI VDFDPQGAATVGLGVNANTVENT  
 IYALFDISVDHVDVHTAFENIDVIPANIDLAAEVOLATVEGQILNSVRLKIK  
 SEYDLIIVDCQPSGLTILNALAADGVIIIPVAEEFFALRGVALLMQSIEKQSRINP  
 ALEVDGLITMTYTKLHCBEVCORIEAFSEKVFHTFISRSIKLPDSTVAAPVYVA  
 PEHTSKREYREARELIARGIVA"  
 4061. -4975  
 /gene="BL1371"  
 4061. -4975  
 /gene="BL1371"  
 /note="COG family: uncharacterized Acr; PFAM\_ID: DUF173"  
 /codon\_start=1  
 /transl\_table=11  
 /product="narrowly conserved hypothetical protein"  
 /protein\_id="AA25172.1"  
 /db\_xref="GI:23326643"  
 /translation="MNETSLRPLPSIREAGSRSEAGNSAEATSTEATSQFQVNL  
 EVYSGPFDALLGMIAANKLELSEVLSISITEFLTVRGDLFTKNWDEASAFDLIASI  
 LRVAKSVAILPGGDSQHQDLSEALREERDLFARLLQYRAYKQAGDFRARTANNG  
 RFPHPAAMDEGIAAMLPELVTLTLELAQAAQVIANAPSEVSHOLHPLVDLRA  
 QSLVVRDLRTAALSKGDSQSFSELTDRCTSRIEVVARFMAVLVFFKQGLVQYQDQ  
 PFALHLRWVSGVDETNDVNISEGDFP"  
 4988. -5704  
 /gene="BL1372"  
 4988. -5704  
 /gene="BL1372"  
 /note="COG family: predicted transcriptional regulator  
 containing the HTH domain"  
 /codon\_start=1  
 /transl\_table=11  
 /product="conserved hypothetical protein"  
 /protein\_id="AA25173.1"  
 /db\_xref="GI:23326644"  
 95. .2941  
 /gene="BL1369"

gene  
CDS  
/translation="MTERIDATNATEHMADAESTAGSAAKAAVRGVEETRPVYDFTVD  
DPFGGLPACLEALMVADOPATATLALILALDRNTVEALQALHDDVEGDSRGRVP  
RPFELRRTAGRQGEASRAAFEPVSVFVDTGQTARLSQALEALAIYVRQVPTAQV  
AAIRGVNSDGVRSITVRGLIKEDGVDPESRAALLVTSFFLDQMGKSLDELPLSLAP  
FLPSVAEMTQSKSGDIDVIS"  
5836..6663  
/gene="BL1373"  
CDS  
5836..6663  
/gene="BL1373"  
/note="COG family: ADP-ribose pyrophosphatase; PFAM\_ID:  
NUDIX"  
/codon\_start=1  
/transl\_table=11  
/product="narrowly conserved hypothetical protein possibly  
in the MutY/Nudix family"  
/protein\_id="AAN25174.1"  
/db\_xref="GI:23326645"

gene  
CDS  
/translation="MGFGNVSERRSAPQGVGVVILALGPKGDTGNSAADSNGAANS  
ASRLWLPLVKKRQPFVFLGCWALPGDLRSLEQSAAYAALESTTDLHPYLEQLVT  
FGSPTRSHGGLPMVSVVYALVQGTAAAGFEDGNVRFPEDELPELAFDHRQIIDVA  
LLRRSKTEYSDVATRLIGPTFTLRQLHGVEAIGGETLDLANFRKMLASGDLDTG  
EKVREGROPATVYRIPKRPVENGTPHESDGDIVELSRQGERDALDALIPSARI"  
6724..8004  
/gene="nadA"  
/note="synonym: BL1374"  
6724..8004  
/gene="nadA"  
/note="nadA; COG family: quinolinate synthase; PFAM\_ID:  
NadA"  
/codon\_start=1  
/transl\_table=11  
/product="quinolinate synthetase"  
/protein\_id="AAN25175.1"  
/db\_xref="GI:23326646"

gene  
CDS  
/translation="MTALSTAPSVDEIIAKLGASTCDAGLTQDPWHFDITTPSYGPG  
ASMLDRPANAPROQVLPDEYRKASDEELQORTADAKQRLGSKLLILGHFYQDEIILK  
HADVQGSFOLAKNATERPADHLVFCGVFMATADILSTPEOSVTLPLNSAGCSMA  
DMANIDVQSCWQGLGICGTGPDSDGLQIIPVTYMNSSAALKAFKGRNGIVCTSS  
NAKALWEARAGKRVLFPPDQHLGRNTARANGIPLSEMLPMDPFKAQGGAADPADYA  
KAMILMGFCSVHQRTFVEQVERAKAYPGVKVIHPECSMQVDAADGTGTATIV  
KEINAPAGSAIAVAGTEINLVRLAAQYPKTVFCLDPPVVCSTWYRHPAYLAWAL  
ENLEAGNVNRTIVDEDTAREAKVALERMLRVHP"  
8092..9723  
/gene="nadB"  
/note="synonym: BL1375"  
8092..9723  
/gene="nadB"  
/note="quinolinate synthetase B; NadB; COG family:  
aspartate oxidase; PFAM\_ID: pyr\_redox; PFAM\_ID:  
FAD\_binding\_2; PFAM\_ID: succ\_DH\_flav\_C"  
/codon\_start=1  
/transl\_table=11  
/product="L-aspartate oxidase"  
/protein\_id="AAN25176.1"  
/db\_xref="GI:23326647"

gene  
CDS  
/translation="MSNTEIVVVGAGVAGLSAALAVAGTGHVTLVKAGLVESNTYH  
AOGGIAAIESDDDPKLHADTWAAGHGLCDPRAVDILTREGAERREFAAAGVHFD  
DAGGMRLERARSRVHVHAGDGTGKVLDDVSAMVRKNPRIHIENAFKDLIV  
RCHTAGVRLIIDGQDKLAAGRVILATGAGRMYPYTNPNPVSTADGLAAARAGAO  
VADLEYQHPHTAMAVGEHFLVSEAVREGCVALLDEHGRYMKDIDPRAELAPRDVVA  
RENFRVMOTQGGKPMVDVSPMTRENPDLASFLAHRFPPTIDATYRSLGFDWSKELIPIV  
APAAHYVMGGIRITLDARTSIPLVYAAGECARTGMGNSRLASNLISLEGLVYGRAGL  
AAVADGDVAWAPATGATGAFDPAIALAAPATTAADDTADKVRDRTRIQNT  
MMHGVGLFDEAGLKTALIAELGGLAANAQAQDINADGAASSVEALENNMLTGVYA  
ATAALARTESRGAHARTIDYSEPPVDNWAHSVAYIKO"  
9727..10620  
/gene="nadC"  
/note="synonym: BL1376"  
9727..10620  
/gene="nadC"  
/note="quinolinate phosphoribosyltransferase  
[decarboxylating]; Caprtase; NadC; COG family:  
nicotinate-nucleotide pyrophosphorylase; PFAM\_ID:

QRPTase N; PFAM\_ID: QRPTase"  
/codon\_start=1  
/transl\_table=11  
/product="probable nicotinate-nucleotide pyrophosphorylase  
[carboxylating]"  
/protein\_id="AAN25177.1"  
/db\_xref="GI:23326648"  
/translation="MLTSHIIRVAVEAALAEADAPYGDITCTTTIPADETGSAHLTARE  
NGVMGIDVFAAFAAQAQNPVITAAIKDGERFORGOILATVEGVPRLLTAEIALLN  
FTQMSGIATWTAFAFVDAVNAIYGDYDGPVTRPRRYERTRIVDTRKTPGLRPFKEY  
GVVCGGHNHYGLSDAVMMKDNHLLAALAAARGIDLAGAIRHVREOVGHTTHIEVEVDR  
LDQIPAVIAGGADTIIMLDNPFSLDITRRKVELIDGKAIVEASGNMSLERVPAVATGVD  
VISVGAUTHSVRSIDLGLDMN"  
10623..11870  
/gene="BL1377"  
10623..11870  
/gene="BL1377"  
/note="COG family: cysteine sulfinate desulfinate/cysteine  
desulfurase and related enzymes; PFAM\_ID: aminotran\_5"  
/codon\_start=1  
/transl\_table=11

Query Match 68.7%; Score 20.6; DB 1; Length 13438;  
Best Local Similarity 85.2%; Pred. No. 1.7e+02;  
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 3 CGGATCCAGCGCCTCTCTTTTGATGGC 29  
|||||  
Db 1327 CGGATCCAGCGCCTCTCTCTCGATGGC 1353  
|||||

Search completed: August 27, 2003, 02:34:23  
Job time : 131.739 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:28:23 ; Search time 11.6853 Seconds  
(without alignments)  
6930.345 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30

Sequence: 1 tccgacccagcgcctctgtttgatggct 30

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	20	AA03472
2	30	100.0	30	21	AA062510
3	28	93.3	1133	20	AA03474
4	28	93.3	1133	21	AA062512
5	20.6	68.7	349980	24	AB081848
6	19.8	66.0	6840	24	AB084742
7	19.8	66.0	6840	24	AB084742
8	19.8	66.0	6843	18	AA084509
					Wolinnella succinog
					W. succinogenes L-
					Wolinnella succinog
					W. succinogenes as
					Bifidobacterium lo
					Human cDNA differe
					Pain regulated cDN
					Human LDL receptor

## ALIGNMENTS

### RESULT 1

AA03472

ID AA03472 standard; DNA; 30 BP.

XX AA03472;

AC AA03472;

XX 30-APR-1999 (first entry)

XX Wolinnella succinogenes L-asparaginase PCR primer 1.

XX L-asparaginase; amplification; treatment; disease; asparagine depletion;

XX malignant disease; haematology; lymphoma; leukaemia; myeloma; AIDS;

XX autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;

XX covalent modification; acylation; pharmacokinetic; immunogenic; spleen;

XX hypersensitive; reduced toxicity; immunosuppression; allergy; thymus;

XX lymphocyte marker distribution; hepatotoxicity; PCR primer; ss.

XX Synthetic.

OS Wolinnella succinogenes.

XX WO9856410-A1.

XX 17-DEC-1998.

XX 09-JUN-1998; 98WO-US11905.

XX 09-JUN-1997; 97US-0049085.

XX (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.

XX Durden DL;

Human late stage o  
Human bladder can  
cDNA encoding huma  
Human benign prost  
Kidney cancer rela  
Human breast cance  
Novel human cDNA s  
Lung cancer relate  
Partial genomic DN  
Novel human polynu  
cDNA encoding nove  
DNA encoding novel  
E. coli DNA for ce  
Drosophila melanog  
Genomic sequence #  
DNA encoding novel  
Drosophila melanog  
Drosophila melanog  
Novel human coding  
Drosophila melanog  
Drosophila melanog  
Aspergillus fumiga  
Aspergillus fumiga  
Rabbit LDL recepto  
Arabidopsis thalia  
Corn ear-derived p  
Human secreted pro  
Human musculoskele  
cDNA encoding nove  
Arabidopsis thalia  
Arabidopsis thalia  
Human cDNA encodin  
Human novel polynu  
Human polynucleoti  
Human polynucleoti  
DNA encoding human

9 19.8 66.0 6902 22 AAF98724  
10 19.8 66.0 6906 25 AAC05992  
11 19.8 66.0 6981 24 ABS76520  
12 19.8 66.0 6981 24 ABK64741  
13 19.8 66.0 6981 24 ABL68557  
14 19.8 66.0 7460 25 ABT31940  
15 19.4 64.7 7308 24 ABL69968  
16 19.4 64.7 4766 19 AAV42563  
17 18.6 62.0 352 22 AAF65999  
18 18.4 61.3 574 23 ABX41664  
19 18.4 61.3 1317 23 AAS88555  
20 18.4 61.3 1365 23 ABL18335  
21 18.4 61.3 1673 23 ABK42208  
22 18.4 61.3 2407 23 AAS79968  
23 18.4 61.3 2688 23 ABL09869  
24 18.4 61.3 5347 23 ABH11058  
25 18.4 61.3 6630 24 ABN59745  
26 18.4 61.3 6835 23 ABL09868  
27 18.4 61.3 7875 23 ABL18334  
28 18.4 61.3 12816 23 ABL20852  
29 18.4 61.3 2621 25 ABT17801  
30 18.2 60.7 2621 25 ABT19615  
31 18.2 60.7 6961 18 AAT84508  
32 18.2 60.7 162 21 AAC50042  
33 18.2 60.7 162 21 ABX84590  
34 18.0 60.0 455 21 AAC31008  
35 18.0 60.0 1333 22 AAL37539  
36 18.0 60.0 1333 25 ABX60827  
37 18.0 60.0 1686 21 AAC42929  
38 18.0 60.0 1686 24 ABZ14235  
39 18.0 60.0 1695 22 AAS26188  
40 18.0 60.0 1695 25 ABX73529  
41 18.0 60.0 2380 24 ABZ11903  
42 18.0 60.0 2452 24 ABZ11904  
43 18.0 60.0 5210 24 ABK52202  
44 18.0 60.0 5210 24 ABK52202  
45 18.0 60.0 5210 24 ABK52202

XX WPI; 1999-070295/06.  
 XX Treating diseases responsive to asparagine depletion - especially  
 PT cancer and auto-immune diseases, using a Wolinella succinogenes  
 PT asparaginase  
 XX  
 PS Example 4; Fig 1; 58pp; English.  
 XX  
 CC This sequence is a PCR primer used in the isolation and amplification of  
 CC an L-asparaginase gene from Wolinella succinogenes which is used in a  
 CC method for the treatment of diseases that respond to asparagine depletion  
 CC This asparaginase and its analogues, are particularly used to treat (i)  
 CC malignant diseases, particularly where haematological and specifically  
 CC lymphoma, leukaemia and myeloma (in either chronic or acute phases) and  
 CC (ii) autoimmune diseases, specifically rheumatoid arthritis, systemic  
 CC lupus erythematosus and acquired immune deficiency syndrome, in humans  
 CC or other mammals. The asparaginase can also be used for the covalent  
 CC modification of biologically active proteins by acylation which is used  
 CC to alter pharmacokinetic and immunogenic properties of biologically  
 CC active proteins, and can be applied to any protein used to treat animal,  
 CC human or plant diseases. The asparaginase can be used to treat patients  
 CC who have become hypersensitive to other microbial asparaginases. The  
 CC asparaginase has reduced toxicity and reduced tendency to cause  
 CC immunosuppression or allergy, particularly it has no effect on spleen and  
 CC thymus histology or lymphocyte marker distribution and no hepatotoxicity.  
 XX  
 SQ Sequence 30 BP; 3 A; 9 C; 8 G; 10 T; 0 other;  
 Query Match 100.0%; Score 30; DB 20; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 0.0006;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCCGGATCCAGCGCCTCTGTTTGTGGCT 30  
 Db 1 TCCGGATCCAGCGCCTCTGTTTGTGGCT 30  
 RESULT 2  
 ID AAC62510  
 AC AAC62510 standard; DNA; 30 BP.  
 XX AAC62510;  
 XX  
 DT 07-FEB-2001 (first entry)  
 XX  
 DE W. succinogenes L-asparaginase PCR primer #1.  
 XX  
 KW Asparaginase; glutaminase; autoimmune disease; graft versus host disease;  
 KW leukaemia; PCR primer; ss.  
 XX  
 OS Wolinella succinogenes.  
 XX  
 PN WO200059533-A1.  
 XX  
 PD 12-OCT-2000.  
 XX  
 PF 24-MAR-2000; 2000WO-US07981.  
 XX  
 PR 02-APR-1999; 99US-0127662.  
 XX  
 PA (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.  
 XX  
 PI Durden DL;  
 XX  
 WPI; 2000-664963/64.  
 XX  
 DR Treating or preventing graft versus host and autoimmune diseases, e.g.  
 PT diabetes, rheumatoid arthritis, that responds to asparagine or  
 PT Glutamine depletion comprises administering an asparaginase or  
 PT glutaminase -  
 XX  
 XX Example 4; Fig 1; 72pp; English.  
 PS

XX The present invention relates to a novel method of using the recombinant  
 CC microbial enzymes asparaginase and glutaminase to treat autoimmune  
 CC diseases and leukaemia. The diseases likely to respond to such treatment  
 CC include graft versus host disease, rheumatoid arthritis, systemic lupus  
 CC erythematosus, autoimmunity, collagen vascular diseases, AIDS,  
 CC osteoarthritis, Isaac's syndrome, psoriasis, diabetes, multiple  
 CC sclerosis, sclerosing panencephalitis, rheumatic fever, inflammatory  
 CC bowel diseases; primary biliary cirrhosis, chronic active hepatitis,  
 CC glomerulonephritis, myasthenia gravis, pemphigus vulgaris and Graves'  
 CC disease. The present sequence is a PCR primer used in the amplification  
 CC of the W. succinogenes asparaginase sequence.  
 XX  
 SQ Sequence 30 BP; 3 A; 9 C; 8 G; 10 T; 0 other;  
 Query Match 100.0%; Score 30; DB 21; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 0.0006;  
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCCGGATCCAGCGCCTCTGTTTGTGGCT 30  
 Db 1 TCCGGATCCAGCGCCTCTGTTTGTGGCT 30  
 RESULT 3  
 ID AAX03474  
 AC AAX03474 standard; DNA; 1133 BP.  
 XX AAX03474;  
 XX  
 DT 30-APR-1999 (first entry)  
 XX  
 DE Wolinella succinogenes L-asparaginase DNA.  
 XX  
 KW L-asparaginase; amplification; treatment; disease; asparagine depletion;  
 KW malignant disease; haematology; lymphoma; leukaemia; myeloma; AIDS;  
 KW autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;  
 KW covalent modification; acylation; pharmacokinetic; immunogenic; spleen;  
 KW hypersensitive; reduced toxicity; immunosuppression; allergy; thymus;  
 KW lymphocyte marker distribution; hepatotoxicity; ss.  
 XX  
 OS Wolinella succinogenes.  
 XX  
 PN WO9856410-A1.  
 XX  
 PD 17-DEC-1998.  
 XX  
 PF 09-JUN-1999; 98WO-US11905.  
 XX  
 PR 09-JUN-1997; 97US-0049085.  
 XX  
 PA (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.  
 XX  
 PI Durden DL;  
 XX  
 WPI; 1999-070295/06.  
 XX  
 DR Treating diseases responsive to asparagine depletion - especially  
 PT cancer and auto-immune diseases, using a Wolinella succinogenes  
 PT asparaginase  
 XX  
 PS Example 5; Fig 6; 58pp; English.  
 XX  
 CC This sequence is a novel L-asparaginase gene from Wolinella succinogenes  
 CC which is used in a method for the treatment of diseases that respond to  
 CC asparagine depletion. This asparaginase and its analogues, are  
 CC particularly used to treat (i) malignant diseases, particularly where  
 CC haematological and specifically lymphoma, leukaemia and myeloma (in  
 CC either chronic or acute phases) and (ii) autoimmune diseases,  
 CC specifically rheumatoid arthritis, systemic lupus erythematosus and  
 CC acquired immune deficiency syndrome, in humans or other mammals. The  
 CC asparaginase can also be used for the covalent modification of  
 CC biologically active proteins by acylation which is used to alter

CC pharmacokinetic and immunogenic properties of biologically active  
 CC proteins, and can be applied to any protein used to treat animal, human  
 CC or plant diseases. The asparaginase can be used to treat patients who  
 CC have become hypersensitive to other microbial asparaginases. The  
 CC asparaginase has reduced toxicity and reduced tendency to cause  
 CC immunosuppression or allergy, particularly it has no effect on spleen and  
 CC thymus histology or lymphocyte marker distribution and no hepatotoxicity.

XX SQ Sequence 1133 BP; 315 A; 305 C; 267 G; 246 T; 0 other;

Query Match 93.3%; Score 28; DB 20; Length 1133;  
 Best Local Similarity 100.0%; Pred. No. 0.0081;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CGGATCCAGCGCTCTGTTTGTGCT 30

DB 96 CGGATCCAGCGCTCTGTTTGTGCT 123

RESULT 4

AC62512  
 ID AAC62512 standard; DNA; 1133 BP.

XX AC AAC62512;

XX DT 07-FEB-2001 (first entry)

XX DE W. succinogenes asparaginase-specific DNA insert.

XX KW Asparaginase; glutaminase; autoimmune disease; graft versus host disease;  
 XX KW leukaemia; DNA insert; ds.

XX OS Wolinella succinogenes.

XX PN WO200059533-A1.

XX PD 12-OCT-2000.

XX PF 24-MAR-2000; 2000WO-US07981.

XX PR 02-APR-1999; 99US-0127662.

XX PA (CHIL-) CHILDRENS HOSPITAL LOS ANGELES.

XX PI Durden DI;

XX PR WPI; 2000-664963/64.

XX PT Treating or preventing graft versus host and autoimmune diseases, e.g.  
 PT diabetes, rheumatoid arthritis, that responds to asparagine or  
 PT glutamine depletion comprises administering an asparaginase or  
 PT glutaminase -

XX PS Example 5; Fig 6; 72pp; English.

XX CC The present invention relates to a novel method of using the recombinant  
 CC microbial enzymes asparaginase and glutaminase to treat autoimmune  
 CC diseases and leukaemia. The diseases likely to respond to such treatment  
 CC include graft versus host disease, rheumatoid arthritis, systemic lupus  
 CC erythematosus, autoimmunity, collagen vascular diseases, AIDS,  
 CC osteoarthritis, Isaac's syndrome, psoriasis, diabetes, multiple  
 CC sclerosis, sclerosing pancreatitis, rheumatic fever, inflammatory  
 CC bowel diseases, primary biliary cirrhosis, chronic active hepatitis,  
 CC glomerulonephritis, myasthenia gravis, pemphigus vulgaris and Graves'  
 CC disease. The present invention is DNA insert used to demonstrate the  
 CC method of the invention.

XX SQ Sequence 1133 BP; 315 A; 305 C; 267 G; 246 T; 0 other;

Query Match 93.3%; Score 28; DB 21; Length 1133;  
 Best Local Similarity 100.0%; Pred. No. 0.0081;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 CGGATCCAGCGCTCTGTTTGTGCT 30

DB 96 CGGATCCAGCGCTCTGTTTGTGCT 123

RESULT 5

ABQ81848  
 ID ABQ81848 standard; DNA; 349980 BP.

XX AC ABQ81848;

XX DT 19-NOV-2002 (first entry)

XX DE Bifidobacterium longum NCC2705 related nucleotide sequence SEQ ID:1104.

XX KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;  
 XX KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;  
 XX KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;  
 XX KW rotavirus; food composition; pharmaceutical composition; gene; ds.

XX OS Bifidobacterium longum.

XX OS Synthetic.

XX PN EPI227152-A1.

XX PD 31-JUL-2002.

XX PF 30-JAN-2001; 2001EP-0102050.

XX PR 30-JAN-2001; 2001EP-0102050.

XX PA (NEST) SOC PROD NESTLE SA.

XX DR WPI; 2002-668397/72.

XX PT Novel polynucleotide comprising Bifidobacterium genome sequence useful  
 PT as a probe or primer for detecting and/or identifying Bifidobacterium  
 PT longum in a biological sample -

XX PS Disclosure; SEQ ID 1104; 80pp; English.

XX CC The present invention describes a polynucleotide (I) comprising a  
 CC sequence of a Bifidobacterium genome selected from the nucleotide  
 CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at  
 CC least 90% identity or which hybridises with the sequences given in  
 CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding  
 CC a fusion protein, comprising a sequence selected from 1097 sequences  
 CC given in ABP65258 to ABP66354 ligated in frame to a polynucleotide  
 CC encoding a heterologous polypeptide. (I) has antidiarrheic and  
 CC antibacterial activities, and can be used as an inhibitor of Salmonella.  
 CC (I) (which is a probe) is useful for the detection and/or identification  
 CC of Bifidobacterium longum in a biological sample. A carrier containing  
 CC the lactic acid bacterium Bifidobacterium longum NCC2705 (NCM I-2618)  
 CC can be used for preventing and/or treating diarrhoea brought about by  
 CC pathogenic bacteria and/or rotavirus. The carrier is a food composition  
 CC selected from milk, yogurt, curd, cheese, fermented milks, milk based  
 CC fermented products, ice-creams, fermented cereal based products, milk  
 CC based powders, infant formula, pet food or a pharmaceutical composition  
 CC selected from tablets, liquid bacterial suspensions, dried oral  
 CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.  
 CC (I) is useful in DNA arrays or chips to carry out analysis of the  
 CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent  
 CC Bifidobacterium related nucleotide sequences given in the Sequence  
 CC Listing from the present invention but not mentioned further within the  
 CC specification.

XX CC N.B. The sequence data for this patent is not represented in the printed  
 CC specification but is based on sequence information supplied by the  
 CC European Patent Office.

XX SQ Sequence 349980 BP; 69195 A; 106952 C; 106128 G; 67705 T; 0 other;

Query Match 68.7%; Score 20.6; DB 24; Length 349980;  
 Best Local Similarity 85.2%; Pred. No. 41;

Matches	23;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

Qy 3 CGATCCAGCGCCTCTGTTTGTATGC 29  
Db 202894 CGATCCAGCGACTCTGCCTCGATGC 202920

## RESULT 6

ABK84742  
ID ABK84742 standard; cDNA; 6840 BP.

AC ABK84742;

DT 14-AUG-2002 (first entry)

DE Human cDNA differentially expressed in granulocytic cells #1313.

KW	Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW	viral infection; parasitic infection; protozoal infection;
KW	fungal infection; sterile inflammatory disease; psoriasis;
KW	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW	cardiac reperfusion injury; renal reperfusion injury; AIDS;
KW	adult respiratory distress syndrome; inflammatory bowel disease;
KW	Crohn's disease; ulcerative colitis; periodontal disease;
KW	granulocyte activation; chronic inflammation; allergy.

OS Homo sapiens.

XX  
PN WO200228999-A2.

11-APR-2002.

03-OCT-2001;

PR 03-OCT-2000; 2000US-237189P.

(GENE-) GENE LOGIC INC.

Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J:

DR WPI; 2002-435328/46.

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity -

PS Claim 1; SEQ ID No 1313; 114pp: English.

The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile

inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and MS is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at <http://wipo.int/pub/published/pct> sequences.

SQ Sequence 6840 BP; 1679 A; 1744 C; 1831 G; 1586 T; 0 other;

Query Match 66.0%; Score 19.8; DB 24; Length 6840;  
Best Local Similarity 91.3%; Pred. No. 53;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28

Db 2514 ATCCAGCGCCTCTGTTGAATGG 2536

## RESULT 7

ABL88419

ID ABL88419 standard; cDNA; 6840 BP.

AC ABL88419;

DT 16-MAY-2002 (first entry)

DE Pain regulated cDNA sequence 62.

KW Pain; analgesic; gene therapy; neurological disorder;

neurodegenerative disease; gene; ss.  
KW

OS Homo sapiens.

PN WO200212338-A2.

14-FEB-2002

03-AUG-2001: 2001WO-EP09011

XX  
PR 03-AUG-2000: 2000DE-1037759XX  
PA (CHEF ) GRIENTHAT. GMRH

XX  
PI Gillen C. Wetzel's T. Wnendt S. Weihe F. Schaefer Mv.

XX  
DR  
WPT: 2002-257469/30

DR P-PSDB; ABB85016.  
XX

PT identifying pain-regulating compounds, useful for treating chronic pain  
PT and for diagnosis, by measuring binding of compounds to specific  
PT peptides and proteins -  
PT

PS Claim 1: Fig 36: 213pp: German.

The invention relates to identifying pain-regulating substances (A) comprises (i) incubating a test substance with a cell (or preparation from it) that has synthesised a peptide or protein (B) and (ii) measuring either binding of the test substance to (B) or some functional parameter that is altered by this binding. The method is useful for identifying pain-regulating substances (A) with analgesic activity. (A) along with nucleic acid (A) that encode proteins (B), AB858006-AB858037) that interact with (A); (B); vectors containing the nucleic acid; antibodies against (B); cells that express (B) and agents that bind to (B), are all useful for treating pain, particularly chronic pain, including use in gene therapy. The same materials can also be used for diagnosis, e.g. of neurological and neurodegenerative diseases. The present sequence is that of a polynucleotide of the invention.

XX SQ Sequence 6840 BP; 1679 A; 1744 C; 1831 G; 1586 T; 0 other;  
 Query Match 66.0%; Score 19.8; DB 24; Length 6840;  
 Best Local Similarity 91.3%; Pred. No. 53;  
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 6 ATCCAGCGCCTCTGTTTGAATGG 28  
 |||||  
 Db 2514 ATCCAGCGCCTCTGTTTGAATGG 2536

RESULT 8  
 AAT84509  
 ID AAT84509 standard; cDNA to mRNA; 6843 BP.  
 XX  
 AC AAT84509;  
 XX  
 DT 01-DEC-1997 (first entry)  
 XX  
 DE Human LDL receptor analogue cDNA.  
 XX  
 DE Low density lipoprotein receptor analogue; LDL receptor;  
 KW lipoprotein; arteriosclerosis; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT CDS  
 FT 81..6725  
 FT Location/Qualifiers  
 FT /\*tag= a  
 FT 81..164  
 FT sig\_peptide  
 FT /\*tag= b  
 FT 165..6722  
 FT mat\_peptide  
 FT /\*tag= c  
 XX  
 PN EP773290-A2.  
 XX  
 PD 14-MAY-1997.  
 XX  
 PF 08-OCT-1996; 96BP-0116108.  
 XX  
 PR 24-APR-1996; 96JP-0102451.  
 PR 09-OCT-1995; 95JP-0261440.  
 XX  
 PA (KOWA ) KOWA CO LTD.  
 XX  
 PI Arai K, Iwasaki A, Saito Y, Yamazaki H;  
 XX  
 WPI; 1997-261317/24.  
 DR P-PSDB; AAW26357.  
 XX  
 PT DNA encoding LDL receptor analogue proteins - for research into  
 PT arteriosclerosis  
 XX  
 PS Claim 3; Page 97-103; 104pp; English.  
 XX  
 CC This full-length cDNA sequence, of which the coding sequence is  
 CC claimed, codes for a novel human low density lipoprotein  
 CC (LDL) receptor family member (AAW26357) which participates in  
 CC lipoprotein metabolism and which is a critical factor that triggers  
 CC the onset of arteriosclerosis. It was identified by screening a  
 CC human brain cDNA library with a partial cDNA clone. A vector  
 CC comprising the coding sequence can be used in a claimed method to  
 CC produce LDL receptor analogue protein in transformant cells. A  
 CC cDNA clone (see AAT84508) coding for a novel rabbit LDL receptor  
 CC (AAW26356) is also claimed. It is expected that, through analysis of  
 CC the protein, details of lipoprotein metabolism mediated by the  
 CC membrane LDL receptor will be elucidated, and the pathology of  
 CC abnormal lipid metabolism which triggers the onset and progression  
 CC of arteriosclerosis will be clarified.  
 XX  
 SQ Sequence 6843 BP; 1682 A; 1744 C; 1831 G; 1586 T; 0 other;

Query Match 66.0%; Score 19.8; DB 18; Length 6843;  
 Best Local Similarity 91.3%; Pred. No. 53;  
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 6 ATCCAGCGCCTCTGTTTGAATGG 28  
 |||||  
 Db 2514 ATCCAGCGCCTCTGTTTGAATGG 2536

RESULT 9  
 AAF98724  
 ID AAF98724 standard; DNA; 6902 BP.  
 XX  
 AC AAF98724;  
 XX  
 DT 02-JUL-2001 (first entry)  
 XX  
 DE Human late stage ovarian tumour polynucleotide marker 39.  
 XX  
 DE Human; ovarian cancer; identification; detection; characterisation;  
 KW tumour; kinase; marker; cytostatic; antisense gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200118542-A2.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 01-SEP-2000; 2000WO-US24199.  
 XX  
 PR 03-SEP-1999; 99US-0152547.  
 PR 16-MAR-2000; 2000US-0190347.  
 PR 21-MAR-2000; 2000US-0191321.  
 PR 31-MAY-2000; 2000US-0208382.  
 PR 20-JUL-2000; 2000US-0220467.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Lee J, Thompsho P, Lillie J;  
 XX  
 WPI; 2001-211428/21.  
 DR  
 PT Detection, assessment, prevention and therapy of ovarian cancer,  
 PT comprises detecting changes in the expression of a variety of markers -  
 XX  
 PS Claim 1; Page 1189-1191; 1199pp; English.  
 XX  
 CC The present invention describes a method for assessing whether a patient  
 CC is afflicted with ovarian cancer by comparing: (1) the expression of a  
 CC marker (I) (see AAF98594 to AAF98730), in a patient sample; and (2) the  
 CC normal level of expression of (I) in a control non-ovarian cancer  
 CC sample, where a significant difference between the level of expression  
 CC in (a) and (b) is an indication that the patient is afflicted with  
 CC ovarian cancer. (I) have cytostatic activities and can be used in  
 CC antisense gene therapy. The method, compositions and kits from the  
 CC present invention can be used for: (1) assessing and treating ovarian  
 CC cancer; (2) making isolated hybridoma, which produces an antibody useful  
 CC for ovarian cancer assessment; and (3) inhibiting ovarian cancer in a  
 CC patient. AAF98573 to AAF98593 represent human kinase marker primers and  
 CC probes which are used in the exemplification of the present invention.  
 XX  
 SQ Sequence 6902 BP; 1719 A; 1747 C; 1837 G; 1599 T; 0 other;

Query Match 66.0%; Score 19.8; DB 22; Length 6902;  
 Best Local Similarity 91.3%; Pred. No. 53;  
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 6 ATCCAGCGCCTCTGTTTGAATGG 28  
 |||||  
 Db 2510 ATCCAGCGCCTCTGTTTGAATGG 2532



Best Local Similarity 91.3%; Pred. No. 53;  
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ATCCAGCGCCTCTGTTTGG 28  
Db 2631 ATCCAGCGCCTCTGTTTGG 2653

RESULT 12  
ABL68557  
ID ABL68557 standard; DNA; 6981 BP.  
XX  
AC ABL68557;  
XX  
DT 18-JUN-2002 (first entry)  
XX  
DE Human benign prostatic hyperplasia gene #636.  
XX  
KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200212440-A2.  
XX  
PD 14-FEB-2002.  
XX  
PF 07-AUG-2001; 2001WO-US24708.  
XX  
PR 07-AUG-2000; 2000US-223323P.  
XX  
PR 05-JUN-2001; 2001US-0873319.  
XX  
PA (GENE-) GENE LOGIC INC.  
PA (NISR) JAPAN TOBACCO INC.  
XX  
PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;  
XX  
DR WPI; 2002-257476/30.  
XX  
XX  
PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by  
PT detecting expression levels of one or more genes in prostate cells from  
PT patient that are differentially regulated compared to normal prostate  
PT cells  
XX  
PS Disclosure; Page 344-346; 444pp; English.  
XX  
CC The invention relates to a method of diagnosing (I) the onset or  
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for  
CC or identifying an agent that modulates the onset or progression of BPH.  
CC The method is based on changes in gene expression in BPH tissue isolated  
CC from patients exhibiting different clinical states of prostate  
CC hyperplasia as compared to normal prostate tissue. (I) comprises  
CC detecting the expression levels of one or more genes in prostate cells  
CC from the subject that are differentially regulated compared to normal  
CC prostate cells. (II) comprises preparing a first gene expression profile  
CC of BPH cells or BPH-like cell population, exposing the cells to the  
CC agent, preparing a second gene expression profile of the agent exposed  
CC cells, and comparing the first and second gene expression profiles.  
CC (I) is useful for diagnosing the onset or progression of BPH. (II) is  
CC useful for identifying an agent that modulates the onset or progression  
CC of BPH. The methods are useful to present information identifying  
CC the expression level in a tissue or cells, by comparing the expression  
CC level of genes given in the specification in the tissue or cells to the  
CC level of expression of gene in the database, and displaying the  
CC expression levels of at least one gene in the tissue or cell sample  
CC compared to the expression level in BPH. Agents using (II) are useful for  
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human  
CC benign prostatic hyperplasia gene sequences of the invention.  
XX  
SQ Sequence 6981 BP; 1720 A; 1771 C; 1866 G; 1624 T; 0 other;

Query Match 66.08; Score 19.8; DB 24; Length 6981;  
Best Local Similarity 91.3%; Pred. No. 53;  
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ATCCAGCGCCTCTGTTTGG 28  
Db 2631 ATCCAGCGCCTCTGTTTGG 2653

RESULT 13  
ABL68557  
ID ABL68557 standard; DNA; 6981 BP.  
XX  
AC ABL68557;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Kidney cancer related gene sequence SEQ ID NO:6894.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235113P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244877P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
PA (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX WPI; 2002-188264/24.  
 XX Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set -  
 XX  
 PS Claim 1; SEQ ID 6894; 44pp; English.  
 XX  
 XX The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 847 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilms' tumour.  
 XX  
 SQ Sequence 6981 BP; 1720 A; 1771 C; 1866 G; 1624 T; 0 other;  
 Query Match 66.0%; Score 19.8; DB 24; Length 6981;  
 Best Local Similarity 91.3%; Pred. No. 53;  
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 6 ATCCAGCGCTCTGTTTGATGG 28  
 DB 2631 ATCCAGCGCTCTGTTTGATGG 2653  
 RESULT 14  
 ABL31940  
 ID ABL31940 standard; DNA; 7460 BP.  
 AC ABL31940;  
 XX  
 XX 01-MAY-2003 (first entry)  
 XX Human breast cancer / ovarian cancer related coding sequence #47.  
 XX Human; gene; ds; cytostatic; breast cancer; ovarian cancer.  
 XX Homo sapiens.  
 XX WO2003000012-A2.  
 XX 03-JAN-2003.  
 XX 21-JUN-2002; 2002WO-US19773.  
 XX 21-JUN-2001; 2001US-300159P.  
 XX 27-JUN-2001; 2001US-301351P.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX Veiby OP;  
 XX WPI; 2003-267848/26.  
 DR P-PSDB; ABJ37071.  
 XX Determining the presence of breast cancer in an individual, involves  
 PT using specific polynucleotide markers -

XX Disclosure; Page 202-204; 233pp; English.  
 PS The invention comprises a method for assessing whether a patient is  
 CC afflicted with breast cancer or ovarian cancer. The method involves the  
 CC use of specific DNA markers. The method of the invention is useful in the  
 CC detection and treatment of ovarian and breast cancer. DNA sequences  
 CC ABL31894 - ABL31949 encode human breast/ovarian cancer-related proteins.  
 XX Sequence 7460 BP; 1844 A; 1890 C; 1992 G; 1734 T; 0 other;  
 SQ  
 Query Match 66.0%; Score 19.8; DB 25; Length 7460;  
 Best Local Similarity 91.3%; Pred. No. 54;  
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 6 ATCCAGCGCTCTGTTTGATGG 28  
 DB 3067 ATCCAGCGCTCTGTTTGATGG 3089  
 RESULT 15  
 ABL70968/C  
 ID ABL70968 standard; cDNA; 570 BP.  
 AC ABL70968;  
 XX 05-MAR-2003 (first entry)  
 XX Novel human cDNA sequence #193.  
 XX Human; gene; ss; nervous system disorder; peripheral neuropathy;  
 KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;  
 KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;  
 KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;  
 KW insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;  
 KW ulcer; burn; bone disorder; osteoporosis; infection; allergic rhinitis; stroke;  
 KW fibrosis; reperfusion injury; osteoarthritis; asthma;  
 KW coagulation disorder; cancer; tumour; inflammatory disease;  
 KW septic shock; Crohn's disease; anaphylaxis; proliferation; chemotactic;  
 KW differentiation; stem cell growth factor; haematopoiesis; chemokinetic;  
 KW haemostatic; antiinflammatory; expressed sequence tag; EST.  
 XX Homo sapiens.  
 XX WO200281731-A2.  
 XX 17-OCT-2002.  
 XX 29-JAN-2002; 2002WO-US01222.  
 XX 30-JAN-2001; 2001US-0774528.  
 XX (HYSE-) HYSEQ INC.  
 XX (GOOD/) GOODRICH R W.  
 XX Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;  
 XX WPI; 2003-058563/05.  
 XX Novel polypeptide useful for treating neurodegenerative diseases,  
 PT myeloid or lymphoid cell disorders, bone disorders, mechanical and  
 PT traumatic disorders, coagulation disorders, and inflammatory diseases  
 XX Claim 1; Page -; 612pp; English.  
 XX This invention relates to the cDNA sequences encoding an isolated  
 CC novel human polypeptide. The protein encoded by the nucleic acid of  
 CC the invention is useful for treating central and peripheral nervous  
 CC system diseases (e.g. peripheral neuropathy, Huntington's disease,  
 CC amyotrophic lateral sclerosis); neurodegenerative diseases (e.g.  
 CC Parkinson's disease, Alzheimer's disease); autoimmune disease (e.g.

CC systemic lupus erythematosus, rheumatoid arthritis, insulin-dependent  
 CC diabetes mellitus); myeloid or lymphoid cell disorders (e.g. anaemia  
 CC and thrombocytopenia); wounds, ulcers, burns; bone disorders (e.g.  
 CC osteoporosis, osteoarthritis); mechanical and traumatic disorders (e.g.  
 CC stroke, head trauma); lung or liver fibrosis; reperfusion injury in  
 CC various tissues; bacterial, viral or fungal infections; allergic  
 CC conditions such as allergic rhinitis, asthma; coagulation disorders  
 CC (e.g. haemophilia); cancer and tumours; and inflammatory diseases (e.g.  
 CC septic shock, Crohn's disease, anaphylaxis). The protein may be used to  
 CC inhibit the growth, infection or function of infectious agents such as  
 CC bacteria, fungi, viruses, or to effect bodily characteristics,  
 CC biorythms or circadian cycles of rhythms. The protein may also  
 CC have proliferation/differentiation, stem cell growth factor,  
 CC haematopoiesis regulation, immune stimulating or suppressing,  
 CC chemotactic/chemokinetic, haemostatic and thrombolytic, receptor/ligand,  
 CC and antiinflammatory activities. The cDNA sequences of the invention are  
 CC useful for expressing recombinant protein for analysis. The present  
 CC sequence represents a novel human cDNA sequence of the invention,  
 CC this sequence is an expressed sequence tag (EST) and was identified  
 CC using subtractive hybridisation.

XX  
 SQ Sequence 570 BP; 151 A; 156 C; 140 G; 123 T; 0 other;

Query Match 64.7%; Score 19.4; DB 25; Length 570;  
 Best Local Similarity 79.3%; Pred. No. 56;  
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTGTTTGATGGCT 30  
 DB 36 CTGGATCCTGAGGCTCTTTTGTGGCT 8

Search completed: August 27, 2003, 01:11:40  
 Job time : 12.6853 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 00:41:09 ; Search time 2.7859 Seconds  
(without alignments)  
4753.042 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30

Sequence: 1 tcggatccaggccctctgtttgatgct 30

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*

2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*

3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*

5: /cgn2\_6/prodata/1/ina/FACTUS COMB.seq:\*

6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	3	US-09-094-435-1
2	28	93.3	1133	3	US-09-094-435-3
3	19.8	66.0	6642	1	US-08-727-034-5
4	19.8	66.0	6843	1	US-08-727-034-6
5	18.2	60.7	6639	1	US-08-727-034-1
6	18.2	60.7	6961	1	US-08-727-034-2
7	18	60.0	277	4	US-09-313-294A-3050
8	17.8	59.3	33	2	US-08-319-795-14
9	17.8	59.3	33	2	US-08-468-985-14
10	17.8	59.3	1180	4	US-09-220-132-169
11	17.8	59.3	2230	3	US-08-448-194-7
12	17.8	59.3	2230	4	US-08-867-921-7
13	17.6	58.7	360	5	PCT-US93-05703-1
14	17.6	58.7	2262	2	US-08-674-887A-5
15	17.6	58.7	2262	3	US-08-951-844-5
16	17.6	58.7	2262	4	US-09-412-347-5
17	17.6	58.7	2469	1	US-07-997-133-2
18	17.6	58.7	2469	1	US-08-459-296-1
19	17.6	58.7	2469	5	US-07-997-133-2
20	17.6	58.7	2662	2	US-08-451-822A-14
21	17.6	58.7	2662	4	US-08-323-430-14
22	17.6	58.7	2733	1	US-08-371-001-14
23	17.6	58.7	2733	5	PCT-US96-00331-14
24	17.6	58.7	4249	1	US-08-480-784-21
25	17.6	58.7	4249	1	US-08-483-553-21
26	17.6	58.7	4249	1	US-08-487-002-21
27	17.6	58.7	4249	1	US-08-483-554B-21

C 28	17.6	58.7	4249	1	US-08-488-011B-21	Sequence 21, Appl
C 29	17.6	58.7	4249	3	US-08-850-727-21	Sequence 21, Appl
C 30	17.6	58.7	4249	5	PCT-US95-10202-21	Sequence 21, Appl
C 31	17.6	58.7	4249	5	PCT-US95-10203-21	Sequence 21, Appl
C 32	17.6	58.7	4249	5	PCT-US95-10220-21	Sequence 21, Appl
C 33	17.6	58.7	5656	1	US-08-425-061-1	Sequence 1, Appl
C 34	17.6	58.7	5656	1	US-08-825-886-1	Sequence 1, Appl
C 35	17.6	58.7	5656	4	US-08-989-890-1	Sequence 3, Appl
C 36	17.6	58.7	5689	1	US-08-425-061-3	Sequence 3, Appl
C 37	17.6	58.7	5689	1	US-08-825-886-3	Sequence 3, Appl
C 38	17.6	58.7	5689	4	US-08-989-890-3	Sequence 11, Appl
C 39	17.6	58.7	5707	1	US-08-425-061-11	Sequence 11, Appl
C 40	17.6	58.7	5707	1	US-08-825-886-11	Sequence 11, Appl
C 41	17.6	58.7	5707	4	US-08-989-890-11	Sequence 2, Appl
C 42	17.6	58.7	5709	1	US-08-425-061-2	Sequence 7, Appl
C 43	17.6	58.7	5709	1	US-08-425-061-7	Sequence 8, Appl
C 44	17.6	58.7	5709	1	US-08-425-061-8	Sequence 9, Appl
C 45	17.6	58.7	5709	1	US-08-425-061-9	

#### ALIGNMENTS

RESULT 1  
US-09-094-435-1  
; Sequence 1, Application US/09094435  
; Patent No. 6251388  
; GENERAL INFORMATION:  
; APPLICANT: Donald L Durden  
; TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES  
; TITLE OF INVENTION: ASPARAGINASE IN THE TREATMENT OF HUMAN  
; TITLE OF INVENTION: HEMATOLOGIC AND AUTOIMMUNE DISEASE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: storage  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: FastSeq for Windows 2.0  
; CURRENT APPLICATION DATA: US/09/094,435  
; APPLICATION NUMBER: US/09/094,435  
; FILING DATE: Filed Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/049,085  
; FILING DATE: June 9, 1997  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 234/274  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 30 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-094-435-1

Query Match 100.0% Score 30; DB 3; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCTCTGTTTGTATGGCT 30  
Db 1 TCCGGATCCAGCGCTCTGTTTGTATGGCT 30

## RESULT 2

US-09-094-435-3  
; Sequence 3, Application US/09094435  
; Patent No. 6251388  
; GENERAL INFORMATION:

APPLICANT: Donald L Durden  
TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES  
TITLE OF INVENTION: ASPARAGINASE IN THE TREATMENT OF HUMAN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage

COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ for Windows 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/094,435  
FILING DATE: Filed Herewith

CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/049,085  
FILING DATE: June 9, 1997  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 234/274  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440

TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1133 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear

US-09-094-435-3

Query Match 93.3%; Score 28; DB 3; Length 1133;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCTCTGTTTGTATGGCT 30  
Db 96 CGGATCCAGCGCTCTGTTTGTATGGCT 123

## RESULT 3

US-08-727-034-5  
; Sequence 5, Application US/08727034  
; Patent No. 5665872  
; GENERAL INFORMATION:

APPLICANT: SAITO, YASHUSHI  
APPLICANT: IWASAKI, AKIO

APPLICANT: ARAI, KOICHI  
APPLICANT: YAMAZAKI, HIROYUKI  
TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND  
TITLE OF INVENTION: THE GENE CODING THEREFOR  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/727,034  
FILING DATE: 08-OCT-1996  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 261440/1995  
FILING DATE: 09-OCT-1995  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 102451/1996  
FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 80-079-0

TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:  
LENGTH: 6642 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA  
US-08-727-034-5

Query Match 66.0%; Score 19.8; DB 1; Length 6642;  
Best Local Similarity 91.3%; Pred. No. 12;  
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCTCTGTTTGTATGG 28  
Db 2434 ATCCAGCGCTCTGTTTGTATGG 2456

## RESULT 4

US-08-727-034-6  
; Sequence 6, Application US/08727034  
; Patent No. 5665872  
; GENERAL INFORMATION:

APPLICANT: SAITO, YASHUSHI  
APPLICANT: IWASAKI, AKIO

APPLICANT: ARAI, KOICHI  
TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND  
TITLE OF INVENTION: THE GENE CODING THEREFOR

NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202

;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/727,034  
;; FILING DATE: 08-OCT-1996  
;; CLASSIFICATION: 536  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 261440/1995  
;; FILING DATE: 09-OCT-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 102451/1996  
;; FILING DATE: 24-APR-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,618  
;; REFERENCE/DOCKET NUMBER: 80-079-0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 6843 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA to mRNA  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 81..6725  
;; OTHER INFORMATION: /note= "Identification Method: S"  
;; FEATURE:  
;; NAME/KEY: sig\_peptide  
;; LOCATION: 81..164  
;; OTHER INFORMATION: /note= "Identification Method: S"  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: 165..6722  
;; OTHER INFORMATION: /function= "Nucleotides 165-6722  
;; OTHER INFORMATION: encode the mature peptide"  
;; OTHER INFORMATION: /note= "Identification Method: S"  
;; US-08-727-034-6

Query Match 66.0%; Score 19.8; DB 1; Length 6843;  
Best Local Similarity 91.3%; Pred. No. 12;  
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ATCCAGCGCCTCTGTTTGATGG 28  
Db 2514 ATCCAGCGCCTCTGTTTGATGG 2536

RESULT 5  
US-08-727-034-1  
; Sequence 1, Application US/08727034  
; Patent No. 5665872  
; GENERAL INFORMATION:  
; APPLICANT: SAITO, YASHUSHI  
; APPLICANT: IWASAKI, AKIO  
; APPLICANT: ARAI, KOICHI  
; APPLICANT: YAMAZAKI, HIROYUKI  
; TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND  
; TITLE OF INVENTION: THE GENE CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA

;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/727,034  
;; FILING DATE: 08-OCT-1996  
;; CLASSIFICATION: 536  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 261440/1995  
;; FILING DATE: 09-OCT-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 102451/1996  
;; FILING DATE: 24-APR-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,618  
;; REFERENCE/DOCKET NUMBER: 80-079-0  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 6639 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA to mRNA  
;; US-08-727-034-1

Query Match 60.7%; Score 18.2; DB 1; Length 6639;  
Best Local Similarity 87.0%; Pred. No. 61;  
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 ATCCAGCGCCTCTGTTTGATGG 28  
Db 2431 ATCCAGCGCCTCTGTTTGATGG 2453

RESULT 6  
US-08-727-034-2  
; Sequence 2, Application US/08727034  
; Patent No. 5665872  
; GENERAL INFORMATION:  
; APPLICANT: SAITO, YASHUSHI  
; APPLICANT: IWASAKI, AKIO  
; APPLICANT: ARAI, KOICHI  
; APPLICANT: YAMAZAKI, HIROYUKI  
; TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND  
; TITLE OF INVENTION: THE GENE CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/727,034  
; FILING DATE: 08-OCT-1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 261440/1995  
; FILING DATE: 09-OCT-1995

; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 102451/1996  
 ; FILING DATE: 24-APR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,618  
 ; REFERENCE/DOCKET NUMBER: 80-079-0  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 6961 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA to mRNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 178..6819  
 ; OTHER INFORMATION: /note= "Identification Method: S"  
 ; FEATURE:  
 ; NAME/KEY: sig peptide  
 ; LOCATION: 178..261  
 ; OTHER INFORMATION: /note= "Identification Method: S"  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: 262..6816  
 ; OTHER INFORMATION: /function= "Nucleotides 262-6816  
 ; OTHER INFORMATION: encode the mature peptide"  
 ; OTHER INFORMATION: /notes= "Identification Method: S"  
 ; US-08-727-034-2

Query Match 60.7%; Score 18.2; DB 1; Length 6961;  
 Best Local Similarity 87.0%; Pred. No. 62;  
 Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28  
 |||||  
 DB 2608 ATCCAGCGCCTCTGTTTGACGG 2630

RESULT 7  
 US-09-313-294A-3050  
 ; Sequence 3050, Application US/09313294A  
 ; Patent No. 6476212  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lalgudi, Raghunath V.  
 ; APPLICANT: Ito, Laura Y.  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
 ; FILE REFERENCE: PL-0017 US  
 ; CURRENT APPLICATION NUMBER: US/09/313,294A  
 ; CURRENT FILING DATE: 1999-05-14  
 ; NUMBER OF SEQ ID NOS: 7600  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 3050  
 ; LENGTH: 277  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Inbyte ID No. 6476212 700610762H1  
 ; NAME/KEY: unsure  
 ; LOCATION: 3, 11, 15, 37, 44-45, 47, 52, 68, 73, 77, 102, 134, 165, 171, 220, 224  
 ; OTHER INFORMATION: a, t, c, g, or other  
 ; US-09-313-294A-3050

Query Match 60.0%; Score 18; DB 4; Length 277;  
 Best Local Similarity 75.0%; Pred. No. 42;  
 Matches 21; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTCTTTTGATGG 28

Db 200 TCGGATCGAGTCATCCGINTTGTGG 227  
 |||||

RESULT 8  
 US-08-319-795-14  
 ; Sequence 14, Application US/08319795  
 ; Patent No. 5980909  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E.  
 ; APPLICANT: Yother, Janet L.  
 ; APPLICANT: McDaniel, Larry S  
 ; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface  
 ; TITLE OF INVENTION: Protein A  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sheomaker and Mattare, Ltd.  
 ; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson  
 ; STREET: Davis Highway  
 ; CITY: Arlington  
 ; STATE: Virginia  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22202-0286  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/319,795  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/246,636  
 ; FILING DATE: 20-MAY-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/048,896  
 ; FILING DATE: 20-APR-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/835,698  
 ; FILING DATE: 12-FEB-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/656,773  
 ; FILING DATE: 15-FEB-1991  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 415-0810  
 ; TELEFAX: (703) 415-0813  
 ; INFORMATION FOR SEQ ID NO: 14:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 33 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-319-795-14

Query Match 59.3%; Score 17.8; DB 2; Length 33;  
 Best Local Similarity 75.9%; Pred. No. 35;  
 Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTCTTTTGATGGCT 30  
 |||||  
 Db 1 CCGGATCCAGCGTCGTATCTTAGGGGCT 29

RESULT 9  
 US-08-468-985-14  
 ; Sequence 14, Application US/08468985  
 ; Patent No. 5997882  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Briles, David E.  
 ; APPLICANT: Yother, Janet L.  
 ; APPLICANT: McDaniel, Larry S  
 ; TITLE OF INVENTION: Epitopic Regions of Pneumococcal Surface

;; TITLE OF INVENTION: Protein A  
;; NUMBER OF SEQUENCES: 20  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sheomaker and Mattare, Ltd.  
;; STREET: 1203 Crystal Plaza Bldg. 1, 2001 Jefferson  
;; STREET: Davis Highway  
;; CITY: Arlington  
;; STATE: Virginia  
;; COUNTRY: U.S.A.  
;; ZIP: 22202-0286  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/468,985  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/319,795  
;; FILING DATE:  
;; APPLICATION NUMBER: US 08/246,636  
;; FILING DATE: 20-MAY-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/048,896  
;; FILING DATE: 20-APR-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/835,698  
;; FILING DATE: 12-FEB-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/656,773  
;; FILING DATE: 15-FEB-1991  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 415-0810  
;; TELEFAX: (703) 415-0813  
;; INFORMATION FOR SEQ ID NO: 14:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 33 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; US-08-468-985-14  
  
Query Match 59.3%; Score 17.8; DB 2; Length 33;  
Best Local Similarity 75.9%; Pred. No. 35;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 2 CCGGATCCAGCGCTCTGTTTGGTGGCT 30  
Db 1 CCGGATCCAGCGCTCTGTTTGGTGGCT 29  
  
RESULT 10  
US-09-220-132-169  
; Sequence 169, Application US/09220132  
; Patent No. 6506607  
; GENERAL INFORMATION:  
; APPLICANT: Stryjan, Andrew W.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
; FILE REFERENCE: 07334-074001  
; CURRENT APPLICATION NUMBER: US/09/220,132  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: US 60/079,303  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: US 60/068,821  
; PRIOR FILING DATE: 1997-12-24  
; NUMBER OF SEQ ID NOS: 191  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 169  
; LENGTH: 1180  
; TYPE: DNA

;; ORGANISM: Homo sapiens  
;; US-09-220-132-169  
  
Query Match 59.3%; Score 17.8; DB 4; Length 1180;  
Best Local Similarity 75.9%; Pred. No. 67;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 2 CCGGATCCAGCGCTCTGTTTGGTGGCT 30  
Db 533 CTGGATCTGGCGCTCTACTTGTAGGGCT 561  
  
RESULT 11  
US-08-448-194-7/c  
; Sequence 7, Application US/08448194  
; Patent No. 6028049  
; GENERAL INFORMATION:  
; APPLICANT: JACOBS, Eric  
; APPLICANT: LEGRAIN, Michele  
; APPLICANT: MAZARIN, Veronique  
; APPLICANT: BOUCHON-THIESEN, Bernadette  
; APPLICANT: SCHRYVERS, Anthony B.  
; APPLICANT: BLOCH, Marie-Aline  
; TITLE OF INVENTION: DNA FRAGMENTS CODING FOR THE TRANSFERRIN  
; TITLE OF INVENTION: RECEPTOR OF NEISSERIA MENINGITIDIS  
; NUMBER OF SEQUENCES: 62  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,194  
; FILING DATE: 23-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/361,469  
; FILING DATE: 22-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/078,053  
; FILING DATE: 18-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 92 07493  
; FILING DATE: 19-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McGowan, Malcolm M.  
; REGISTRATION NUMBER: 39,300  
; REFERENCE/DOCKET NUMBER: 016100-005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2230 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: DNA which encodes Tbp2 subunit of transferrin  
; ORGANISM: receptor  
; STRAIN: Neisseria meningitidis IM2169  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 60..119  
; FEATURE:

NAME/KEY: mat\_peptide  
LOCATION: 120..2192  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 60..2192  
US-08-448-194-7

Query Match 59.3%; Score 17.8; DB 3; Length 2230;  
Best Local Similarity 75.9%; Pred. No. 75;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 2 CCGATCCAGCGCTCTCTTTTGATGGCT 30  
Db 472 CCGTTCAGCGCTCGCGTTTGTGATGTT 444

## RESULT 12

US-08-867-921-7/c  
Sequence 7, Application US/08967921  
Patent No. 6326350  
GENERAL INFORMATION:  
APPLICANT: JACOBS, Eric  
APPLICANT: LEGRAND, Michele  
APPLICANT: MAZARIN, Veronique  
APPLICANT: BOUCHON-THIESEN, Bernadette  
APPLICANT: SCHRYVERS, Anthony B.  
TITLE OF INVENTION: DNA FRAGMENT'S CODING FOR THE TRANSFERRIN  
RECEPTOR OF NEISSERIA MENINGITIDIS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/867,921  
FILING DATE: 03-JUN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/445,472  
FILING DATE: 22-MAY-1995  
APPLICATION NUMBER: US 08/361,469  
FILING DATE: 22-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/078,053  
FILING DATE: 18-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 92 07493  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 016100-004  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2230 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:  
ORGANISM: DNA which encodes Tbp2 subunit of transferrin  
ORGANISM: receptor  
STRAIN: Neisseria meningitidis IM2169  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 60..119  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 120..2192  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 60..2192  
US-08-867-921-7

Query Match 59.3%; Score 17.8; DB 4; Length 2230;  
Best Local Similarity 75.9%; Pred. No. 75;  
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 CCGATCCAGCGCTCTCTTTTGATGGCT 30  
Db 472 CCGTTCAGCGCTCGCGTTTGTGATGTT 444

## RESULT 13

PCT-US93-05703-1/c  
Sequence 1, Application PC/TUS9305703  
GENERAL INFORMATION:  
APPLICANT: Nova, Michael P.  
APPLICANT: Gonzalez, Ana-Maria  
APPLICANT: Baird, J. Andrew  
TITLE OF INVENTION: Process for Detection of Neoplastic  
TITLE OF INVENTION: Disease  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fitch, Even, Tabin & Flannery  
STREET: 135 South LaSalle Street, Suite 900  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/05703  
FILING DATE: 19930614  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/900,646  
FILING DATE: 18-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 54625PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-552-1311  
TELEFAX: 619-552-0095  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 360 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..360  
OTHER INFORMATION: /standard\_name= "flg probe"





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 02:34:35 ; Search time 22.3388 Seconds  
(without alignments)  
3081.078 Million cell updates/sec

Title: US-09-937-982-1

Perfect score: 30

Sequence: 1 tccggatccagcgcctctgtttgatgctt 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09D\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	10	US-09-773-260-1
2	28	93.3	1133	10	US-09-773-260-3
3	19.8	66.0	6868	11	US-09-919-039-39
4	19.8	66.0	6981	12	US-09-873-319-636
5	19.8	66.0	6981	12	US-09-960-706-977
6	19.8	66.0	6981	14	US-10-097-340-299
7	19.8	66.0	7460	10	US-10-176-847-93
8	19.4	64.7	73308	10	US-09-954-456-276
9	19.4	64.7	3186778	13	US-10-027-632-174961
10	18.8	62.7	1010	13	US-10-027-632-255747
11	18.6	62.0	759	13	US-10-027-632-18853
12	18.4	61.3	374	13	US-10-027-632-83004
13	18.4	61.3	378	13	US-10-027-632-38490
14	18.4	61.3	378	13	US-10-027-632-38491
15	18.4	61.3	378	13	US-10-027-632-83005
16	18.4	61.3	574	10	US-09-764-847-62

C 17	18.4	61.3	574	14	US-10-092-154-62	Sequence 62, Appl
C 18	18.4	61.3	1365	9	US-09-815-242-6300	Sequence 6300, Ap
C 19	18.4	61.3	2407	10	US-09-764-847-1095	Sequence 1095, Ap
C 20	18.4	61.3	2407	14	US-10-092-154-1095	Sequence 1095, Ap
C 21	18.4	61.3	12588	11	US-09-992-009-2	Sequence 2, Appl
C 22	18.2	60.7	832	13	US-10-027-632-168244	Sequence 168244,
C 23	18.2	60.7	1173	13	US-10-027-632-251412	Sequence 251412,
C 24	18.2	60.7	2621	14	US-10-128-714-159	Sequence 159, App
C 25	18.2	60.7	2621	14	US-10-128-714-159	Sequence 159, App
C 26	18	60.0	488	11	US-09-918-998-21292	Sequence 21292, A
C 27	18	60.0	854	13	US-10-027-632-173748	Sequence 173748,
C 28	18	60.0	865	13	US-10-027-632-172306	Sequence 172306,
C 29	18	60.0	865	13	US-10-027-632-172307	Sequence 172307,
C 30	18	60.0	1333	10	US-09-764-877-3904	Sequence 3904, Ap
C 31	18	60.0	1686	10	US-09-938-842A-2040	Sequence 2040, Ap
C 32	18	60.0	1695	10	US-09-764-864-367	Sequence 367, App
C 33	17.8	59.3	248	10	US-09-998-598-609	Sequence 609, App
C 34	17.8	59.3	281	10	US-09-960-352-7003	Sequence 7003, Ap
C 35	17.8	59.3	395	10	US-09-960-352-12400	Sequence 12400, A
C 36	17.8	59.3	399	9	US-09-905-243-55	Sequence 55, Appl
C 37	17.8	59.3	443	10	US-09-960-352-5084	Sequence 5084, Ap
C 38	17.8	59.3	447	9	US-09-815-343-829	Sequence 829, App
C 39	17.8	59.3	447	10	US-09-960-352-7936	Sequence 7936, Ap
C 40	17.8	59.3	477	10	US-09-998-598-289	Sequence 289, App
C 41	17.8	59.3	544	10	US-09-920-300A-388	Sequence 388, App
C 42	17.8	59.3	544	13	US-10-033-528-388	Sequence 388, App
C 43	17.8	59.3	550	13	US-10-027-632-185375	Sequence 185375,
C 44	17.8	59.3	585	13	US-10-027-632-226737	Sequence 226737,
C 45	17.8	59.3	585	13	US-10-027-632-226738	Sequence 226738,

## ALIGNMENTS

RESULT 1  
US-09-773-260-1  
; Sequence 1, Application US/09773260  
; Patent No. US20020102251A1

GENERAL INFORMATION:

APPLICANT: Donald L Durden

TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES  
ASPARAGINASE IN THE TREATMENT OF HUMAN  
HEMATOLOGIC AND AUTOIMMUNE DISEASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq for Windows 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/773,260

FILING DATE: 31-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/094,435

FILING DATE: 1998-06-09

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 234/274

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-773-260-1

Query Match 100.0%; Score 30; DB 10; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.00033;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCGGATCCAGCGCCTCTGTTTGATGGCT 30

Db 1 TCCGGATCCAGCGCCTCTGTTTGATGGCT 30

RESULT 2

US-09-773-260-3

Sequence 3, Application US/09773260

Patent No. US20020102251A1

GENERAL INFORMATION:

APPLICANT: Donald L Durden

TITLE OF INVENTION: UTILIZATION OF WOLINELLA SUCCINOGENES

ASPARAGINASE IN THE TREATMENT OF HUMAN

HEMATOLOGIC AND AUTOIMMUNE DISEASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Suite 4700

STATE: Los Angeles

COUNTRY: California

ZIP: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq for Windows 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/773,260

FILING DATE: 31-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/094,435

FILING DATE: 1998-06-09

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 234/274

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1133 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-773-260-3

Query Match 93.3%; Score 28; DB 10; Length 1133;

Best Local Similarity 100.0%; Pred. No. 0.0038;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CGGATCCAGCGCCTCTGTTTGATGGCT 30

Db 96 CGGATCCAGCGCCTCTGTTTGATGGCT 123

RESULT 3

US-09-919-039-39

Sequence 39, Application US/09919039

Publication No. US20030108871A1

GENERAL INFORMATION:

APPLICANT: Kaser, Matthew R.

TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES

FILE REFERENCE: PA-0035 US

CURRENT APPLICATION NUMBER: US/09/919,039

CURRENT FILING DATE: 2002-09-09

PRIOR APPLICATION NUMBER: 60/222,113

PRIOR FILING DATE: 2000-07-28

NUMBER OF SEQ ID NOS: 401

SOFTWARE: PERL Program

SEQ ID NO 39

LENGTH: 6868

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US20030108871A1 1269631CB1

US-09-919-039-39

Query Match 66.0%; Score 19.8; DB 11; Length 6868;

Best Local Similarity 91.3%; Pred. No. 25;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28

Db 2514 ATCCAGCGCCTCTGTTTGATGG 2536

RESULT 4

US-09-873-319-636

Sequence 636, Application US/09873319A

Publication No. US20030134324A1

GENERAL INFORMATION:

APPLICANT: Munger, William E.

APPLICANT: Kulkarni, Prakash

APPLICANT: Getzenberg, Robert H.

APPLICANT: Waga, Iwao

APPLICANT: Yamamoto, Jun

TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic

Hyperplasia Using Gene Expression Profiles

FILE REFERENCE: 44921-5029-US

CURRENT APPLICATION NUMBER: US/09/873,319A

CURRENT FILING DATE: 2001-06-05

EARLIER APPLICATION NUMBER: US 60/223,323

EARLIER FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 755

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 636

LENGTH: 6981

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20030134324A1 U60975

US-09-873-319-636

Query Match 66.0%; Score 19.8; DB 12; Length 6981;

Best Local Similarity 91.3%; Pred. No. 25;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28

Db 2631 ATCCAGCGCCTCTGTTTGATGG 2653

```

RESULT 5
US-09-960-706-977
; Sequence 977, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 977
; LENGTH: 6981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 U60975
US-09-960-706-977

Query Match          66.0%; Score 19.8; DB 12; Length 6981;
Best Local Similarity 91.3%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGTATGG 28
|||||
DB 2631 ATCCAGCGCCTCTGTTTGTATGG 2653

RESULT 6
US-10-097-340-299
; Sequence 299, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen ZHANG
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26

```

```

; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 299
; LENGTH: 6981
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-299

Query Match          66.0%; Score 19.8; DB 14; Length 6981;
Best Local Similarity 91.3%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGTATGG 28
|||||
DB 2631 ATCCAGCGCCTCTGTTTGTATGG 2653

RESULT 7
US-10-176-847-93
; Sequence 93, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 7460
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-847-93

Query Match          66.0%; Score 19.8; DB 14; Length 7460;
Best Local Similarity 91.3%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGTATGG 28
|||||
DB 3067 ATCCAGCGCCTCTGTTTGTATGG 3089

RESULT 8
US-09-954-456-2276/c
; Sequence 2276, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27

```

RESULT 10  
US-10-027-632-255747/c  
; Sequence 255747, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.

```

Query Match      62.0%; Score 18.6; DB 13; Length 759;
Best Local Similarity 77.8%; Pred. No. 72;
Matches 21; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Ov 4 GGATCCAGCGCCTCTGTTTGGTGGCT 30

```

```
Db      248 GGCYCCAGCAGCTCTCTGTTGGCT 274
      || : ||||| ||||| ||||| |||||
RESULT 12
US-10-027-632-83004/c
; Sequence 83004, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83004
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-83004

Query Match      61.3%; Score 18.4; DB 13; Length 374;
Best Local Similarity 78.6%; Pred. No. 84;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 CCGATCCAGCGCTCTGTTGATGCG 29
      ||||| ||||| ||||| |||||
Db      344 CTGGATCCAGATCTCTGTTAGTGCG 317

RESULT 13
US-10-027-632-38490/c
; Sequence 38490, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38490
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-38490

Query Match      61.3%; Score 18.4; DB 13; Length 378;
Best Local Similarity 78.6%; Pred. No. 84;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 CCGATCCAGCGCTCTGTTGATGCG 29
      ||||| ||||| ||||| |||||
Db      348 CTGGATCCAGATCTCTGTTAGTGCG 321

RESULT 14
US-10-027-632-38491/c
; Sequence 38491, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38491
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-38491

Query Match      61.3%; Score 18.4; DB 13; Length 378;
Best Local Similarity 78.6%; Pred. No. 84;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 CCGATCCAGCGCTCTGTTGATGCG 29
      ||||| ||||| ||||| |||||
Db      348 CTGGATCCAGATCTCTGTTAGTGCG 321

RESULT 15
US-10-027-632-83005/c
; Sequence 83005, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38490
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-83005
```

```
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83005
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-83005
```

```
Query Match      61.3%; Score 18.4; DB 13; Length 378;
Best Local Similarity 78.6%; Pred. No. 84;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      2 CCGGATCCAGCGCCTCTGTTTGGTGGC 29
      | | | | | | | | | | | | | | | |
Db      348 CTGGATCCAGAAATCTCTGTTTAGGTGGC 321
```

Search completed: August 27, 2003, 11:17:53  
Job time : 28.388 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 27, 2003, 00:35:29 ; Search time 92.0378 Seconds  
(without alignments)  
7922.115 Million cell updates/sec

Title: US-09-937-982-1  
Perfect score: 30  
Sequence: 1 tccgattccagcgcctctgtttgatggct 30

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estcom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.6	75.3	789	28	BH249624
C 2	21	70.0	197	12	BP016090
3	21	70.0	530	11	CNS09M4Z
C 4	21	70.0	596	13	BW279822

C 5	21	70.0	601	9	AV960848
C 6	21	70.0	644	9	AV674160
C 7	21	70.0	661	9	AV991875
C 8	21	70.0	699	13	BW246571
C 9	21	70.0	758	13	BW344351
10	20.6	68.7	699	9	AU234665
11	20.6	68.7	924	13	BQ931065
12	20	66.7	454	28	AZ557854
13	20	66.7	456	10	BF469784
14	19.8	66.0	215	14	D59983
C 15	19.8	66.0	300	14	C14671
16	19.8	66.0	623	9	AL042811
17	19.8	66.0	882	9	AU132200
18	19.6	65.3	528	9	AA823562
19	19.6	65.3	589	9	AI907267
C 20	19.6	65.3	786	28	BH651954
C 21	19.6	65.3	888	13	BU595186
C 22	19.6	65.3	1592	13	BH246061
23	19.6	65.3	4293	11	AK044094
C 24	19.4	64.7	198	10	BG220871
25	19.4	64.7	336	9	AW356036
26	19.4	64.7	345	9	AW427901
27	19.4	64.7	378	9	AA416116
28	19.4	64.7	435	10	BF601890
29	19.4	64.7	465	9	AA590627
30	19.4	64.7	468	9	AA162206
C 31	19.4	64.7	582	10	BE074228
C 32	19.4	64.7	614	28	AZ915332
C 33	19.4	64.7	618	28	AQ397708
C 34	19.4	64.7	637	29	AG130724
35	19.4	64.7	786	10	BF698692
36	19.4	64.7	896	14	CB198817
C 37	19.4	64.7	914	10	BE789678
C 38	19.4	64.7	976	29	CNS028NY
39	19.4	64.7	1046	13	EX428164
40	19.4	64.7	1175	9	AL515901
41	19.2	64.0	332	14	CA677807
C 42	19.2	64.0	586	29	BZ887403
C 43	19.2	64.0	892	10	BF133459
C 44	19.2	64.0	925	29	CNS033AH
C 45	19.2	64.0	932	29	CNS020SK

## ALIGNMENTS

RESULT 1  
BH249624

LOCUS

DEFINITION

BOGAQ68TR BOGA

Brassica oleracea

survey sequence.

ACCESSION

BH249624

VERSION

BH249624.1

KEYWORDS

GSS.

SOURCE

Brassica oleracea

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BH249624 789 bp DNA linear GSS 26-NOV-2001  
BOGAQ68TR BOGA Brassica oleracea genomic clone BOGAQ68, genomic survey sequence.

BH249624.1 GI:17072437

BH249624

GSS.

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 789)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

Other GSSs: BOGAQ68TF

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: Sheared ends.

```

FEATURES
  source
    Location/Qualifiers
      1..789
        /organism="Brassica oleracea"
        /mol_type="genomic DNA"
        /strain="T01000DH3"
        /db_xref="taxon:3712"
        /clone_lib="BOGAQ88"
        /note="Vector: pHD1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHD1 using BstXI linkers"
BASE COUNT      251 a 130 c 156 g 252 t
ORIGIN
Query Match      75.3%; Score 22.6; DB 28; Length 789;
Best Local Similarity 86.2%; Pred. No. 42;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CCGGATCCAGCGCTCTGTTTGATGGCT 30
Db 512 CCGGAGCCAGCGCTCTGTTTGTTGCT 540

RESULT 2
BP016090/c
LOCUS
DEFINITION
BP016090 Nori Satoh unpublished cDNA library, young adult Ciona
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Ciona intestinalis
  Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
  Phlebobranchia; Cionidae; Ciona.
REFERENCE
  1 (bases 1 to 197)
  Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.
  Expressed genes in Ciona intestinalis
  Unpublished
  Contact: Nori Satoh
  Department of Zoology
  Kyoto University
  Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
  Tel: 81-75-753-4081
  Fax: 81-75-705-1113
  Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES
  source
    Location/Qualifiers
      1..197
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="ciad59e6"
        /tissue_type="whole animal"
        /dev_stage="young adult"
        /clone_lib="Nori Satoh unpublished cDNA library, young adult"
BASE COUNT      69 a 45 c 45 g 37 t 1 others
ORIGIN
Query Match      70.0%; Score 21; DB 12; Length 197;
Best Local Similarity 82.8%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 CCGGATCCAGCGCTCTGTTTGATGGCT 30
Db 74 CTGGGTCCAGGGCTCTGTTTCATGGCT 46

RESULT 3
CNS09M4Z
LOCUS
DEFINITION
Single read from an extremity of a full-length cDNA clone made from Anopheles gambiae total adult females. 3-PRIME end of clone FK0AC48BD03 of strain 6-9 of Anopheles gambiae (African malaria mosquito).
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Anopheles gambiae (African malaria mosquito)
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
  Anopheles.
  1 (bases 1 to 530)
  Genoscope.
  Direct Submission
  Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Location/Qualifiers
    1..530
      /organism="Anopheles gambiae"
      /mol_type="mRNA"
      /strain="6-9"
      /db_xref="taxon:7165"
      /clone="FK0AC48BD03"
      /plasmid="pME18S-FL"
      /note="end : 3-PRIME"
BASE COUNT      99 a 161 c 162 g 108 t
ORIGIN
Query Match      70.0%; Score 21; DB 11; Length 530;
Best Local Similarity 82.8%; Pred. No. 1.8e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TCCGGATCCAGCGCTCTGTTTGATGGC 29
Db 415 TCCGGATCGAGCACCTCTTTTGACGGC 443

RESULT 4
BW279822/c
LOCUS
DEFINITION
BW279822 Nori Satoh unpublished cDNA library, gonad Ciona
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Ciona intestinalis
  Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
  Phlebobranchia; Cionidae; Ciona.
REFERENCE
  1 (bases 1 to 596)
  Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N.
  Expressed genes in Ciona intestinalis (2002c)
  Unpublished
  Contact: Nori Satoh
  Department of Zoology
  Kyoto University
  Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
  Tel: 81-75-753-4081
  Fax: 81-75-705-1113
  Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES
  source
    Location/Qualifiers
      1..596
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="ciG012b14"
        /tissue_type="gonad"
        /clone_lib="Nori Satoh unpublished cDNA library, gonad"
BASE COUNT      200 a 103 c 89 g 201 t 3 others
ORIGIN
Query Match      70.0%; Score 21; DB 13; Length 596;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

RESULT 5
CNS09M4Z
LOCUS
DEFINITION
Single read from an extremity of a full-length cDNA clone made from Anopheles gambiae total adult females. 3-PRIME end of clone FK0AC48BD03 of strain 6-9 of Anopheles gambiae (African malaria mosquito).
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Anopheles gambiae (African malaria mosquito)
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
  Anopheles.
  1 (bases 1 to 530)
  Genoscope.
  Direct Submission
  Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Location/Qualifiers
    1..530
      /organism="Anopheles gambiae"
      /mol_type="mRNA"
      /strain="6-9"
      /db_xref="taxon:7165"
      /clone="FK0AC48BD03"
      /plasmid="pME18S-FL"
      /note="end : 3-PRIME"
BASE COUNT      99 a 161 c 162 g 108 t
ORIGIN
Query Match      70.0%; Score 21; DB 11; Length 530;
Best Local Similarity 82.8%; Pred. No. 1.8e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TCCGGATCCAGCGCTCTGTTTGATGGC 29
Db 415 TCCGGATCGAGCACCTCTTTTGACGGC 443

RESULT 6
BW279822/c
LOCUS
DEFINITION
BW279822 Nori Satoh unpublished cDNA library, gonad Ciona
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Ciona intestinalis
  Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
  Phlebobranchia; Cionidae; Ciona.
REFERENCE
  1 (bases 1 to 596)
  Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N.
  Expressed genes in Ciona intestinalis (2002c)
  Unpublished
  Contact: Nori Satoh
  Department of Zoology
  Kyoto University
  Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
  Tel: 81-75-753-4081
  Fax: 81-75-705-1113
  Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES
  source
    Location/Qualifiers
      1..596
        /organism="Ciona intestinalis"
        /mol_type="mRNA"
        /db_xref="taxon:7719"
        /clone="ciG012b14"
        /tissue_type="gonad"
        /clone_lib="Nori Satoh unpublished cDNA library, gonad"
BASE COUNT      200 a 103 c 89 g 201 t 3 others
ORIGIN
Query Match      70.0%; Score 21; DB 13; Length 596;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

RESULT 7
CNS09M4Z
LOCUS
DEFINITION
Single read from an extremity of a full-length cDNA clone made from Anopheles gambiae total adult females. 3-PRIME end of clone FK0AC48BD03 of strain 6-9 of Anopheles gambiae (African malaria mosquito).
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Anopheles gambiae (African malaria mosquito)
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
  Anopheles.
  1 (bases 1 to 530)
  Genoscope.
  Direct Submission
  Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Location/Qualifiers
    1..530
      /organism="Anopheles gambiae"
      /mol_type="mRNA"
      /strain="6-9"
      /db_xref="taxon:7165"
      /clone="FK0AC48BD03"
      /plasmid="pME18S-FL"
      /note="end : 3-PRIME"
BASE COUNT      99 a 161 c 162 g 108 t
ORIGIN
Query Match      70.0%; Score 21; DB 11; Length 530;
Best Local Similarity 82.8%; Pred. No. 1.8e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TCCGGATCCAGCGCTCTGTTTGATGGC 29
Db 415 TCCGGATCGAGCACCTCTTTTGACGGC 443

```

```

QY 2 CCGGATCCAGCGCCTCTGTTTGTGGCT 30
Db 72 CTGGGTCCAGGGCTTCTGTTTTCATGGCT 44

RESULT 5
AV960848/c
LOCUS AV960848 Nori Satoh unpublished cDNA library, cleavage stage embryo
DEFINITION Ciona intestinalis cDNA clone cici14b08 5', mRNA sequence.
ACCESSION AV960848
VERSION AV960848.1 GI:19449147
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 601)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1..601
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cici14b08"
/tissue_type="whole animal"
/dev_stage="cleavage stage embryo"
/clone_lib="Nori Satoh unpublished cDNA library, cleavage
stage embryo"
BASE COUNT 176 a 129 c 134 g 162 t
ORIGIN
Query Match 70.0%; Score 21; DB 9; Length 601;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTGTTTGTGGCT 30
Db 580 CTGGGTCCAGGGCTTCTGTTTTCATGGCT 552

RESULT 6
AV674160/c
LOCUS AV674160 Nori Satoh unpublished cDNA library Ciona intestinalis
DEFINITION Ciona clone cibt14j14 5', mRNA sequence.
ACCESSION AV674160
VERSION AV674160.1 GI:10112159
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 644)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113

QY 2 CCGGATCCAGCGCCTCTGTTTGTGGCT 30
Db 580 CTGGGTCCAGGGCTTCTGTTTTCATGGCT 552

RESULT 7
AV991875/c
LOCUS AV991875 Nori Satoh unpublished cDNA library, cleavage stage embryo
DEFINITION Ciona intestinalis cDNA clone cici157d16 5', mRNA sequence.
ACCESSION AV991875
VERSION AV991875.1 GI:19480357
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 661)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoheascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1..661
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cici157d16"
/tissue_type="whole animal"
/dev_stage="cleavage stage embryo"
/clone_lib="Nori Satoh unpublished cDNA library, cleavage
stage embryo"
BASE COUNT 201 a 135 c 143 g 181 t 1 others
ORIGIN
Query Match 70.0%; Score 21; DB 9; Length 661;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTGTTTGTGGCT 30
Db 620 CTGGGTCCAGGGCTTCTGTTTTCATGGCT 592

RESULT 8
BW246571/c
LOCUS BW246571 Nori Satoh unpublished cDNA library, tailbud embryo Ciona
DEFINITION intestinalis cDNA clone cibt07ic16 5', mRNA sequence.
ACCESSION BW246571

```

```

VERSION      BW246571.1  GI:24826489
KEYWORDS     EST.
SOURCE       Ciona intestinalis
ORGANISM     Ciona intestinalis
             Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
             Phlebobranchia; Clonidae; Ciona.
REFERENCE    1 (bases 1 to 699)
AUTHORS      Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE        Expressed genes in Ciona intestinalis (2002c)
JOURNAL      Unpublished
COMMENT      Contact: Nori Satoh
             Department of Zoology
             Kyoto University
             Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
             Tel: 81-75-753-4081
             Fax: 81-75-705-1113
             Email: satoheascidian.zool.kyoto-u.ac.jp.
FEATURES     source
             Location/Qualifiers
               1..699
               /organism="Ciona intestinalis"
               /mol_type="mRNA"
               /db_xref="taxon:7719"
               /clone="citb071c16"
               /tissue_type="whole animal"
               /dev_stage="tailbud embryo"
               /clone_lib="Nori Satoh unpublished cDNA library, tailbud
               embryo"
BASE COUNT   211 a 145 c 158 g 185 t
ORIGIN
Query Match 70.0%; Score 21; DB 13; Length 699;
Best Local Similarity 82.8%; Pred. No. 2e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 CCGGATCCAGCGCCTCTGTTTGATGGCT 30
Db 636 CTGGTCCAGGGCTCTGTTTCATGGCT 608

RESULT 9
LOCUS      BU344351
DEFINITION BU344351
ACCESSION BU344351.1  GI:25852352
KEYWORDS   EST.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
             Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 758)
AUTHORS    Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
             Fong,W.R., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE      A Comprehensive Collection of Chicken cDNAs
JOURNAL    Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE    22335534
PUBMED     12445392
COMMENT    Contact: Simon Hubbard
             Department of Biomolecular Sciences
             University of Manchester Institute of Science and Technology (UMIST)
             )
             PO Box 88, Manchester, M60 1QD, UK
             Tel: 01612008930
             Fax: 01612360409
             Email: Simon.Hubbard@umist.ac.uk.
             Location/Qualifiers
               1..758
               /organism="Gallus gallus"
               /mol_type="mRNA"
               /strain="Compton Line 151"
               /db_xref="taxon:9031"
FEATURES     source
             Location/Qualifiers
               1..699
               /organism="Branchiostoma belcheri"
               /mol_type="mRNA"
               /db_xref="taxon:7741"
               /clone="Bb5_02D02"
               /tissue_type="notochord"
               /clone_lib="Amphioxus Notochord cDNA library"
BASE COUNT   181 a 159 c 211 g 148 t
ORIGIN
Query Match 68.7%; Score 20.6; DB 9; Length 699;
Best Local Similarity 85.2%; Pred. No. 2.9e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 GGATCCAGCGCCTCTGTTTGATGGCT 30

```

JOURNAL  
COMMENT

Db 556 GGGTCAGCGCTACTGTTGATGGCT 582  
|||||

## RESULT 11

## BQ931065

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

BQ931065 8993468 NCI CGAP Co24 Mus musculus cDNA clone  
IMAGE:6399431 5', mRNA sequence.

AGENCOURT 8993468 NCI CGAP Co24 Mus musculus cDNA clone  
IMAGE:6399431 5', mRNA sequence.

EST. 924 bp mRNA linear EST 20-AUG-2002  
Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 924)  
NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished

Contact: Robert Strausberg, Ph.D.  
Email: cgaphs@mail.nih.gov

Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov  
Plate: LLNL3900 row: h column: 24  
High quality sequence stop: 432.

Location/Qualifiers  
1. .924  
/organism="Mus musculus"  
/mol\_type="mRNA"

/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6399431"  
/lab\_host="DH10B (T1 phage-resistant)"

/notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 185 a 288 c 251 g 199 t  
ORIGIN

Query Match 68.7%; Score 20.6; DB 13; Length 924;  
Best Local Similarity 85.2%; Pred. No. 3.1e+02;  
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 GGATCCAGCGCCTCTGTTTGATGGCT 30  
Db 616 GGCTCTGTCCTCTGTTTGATGGCT 642

## RESULT 12

## AZ557854

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

AZ557854 454 bp DNA linear GSS 20-NOV-2000  
RPCI-23-205J5-TV RPCI-23 Mus musculus genomic clone RPCI-23-205J5,  
genomic survey sequence.

AGENCOURT 454 bp DNA linear GSS 20-NOV-2000  
IMAGE:6399431 5', mRNA sequence.

EST. 454 bp DNA linear GSS 20-NOV-2000  
Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 454)  
Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatsman, S., Akinret  
B., Levens, M., McGann, S., Isegaye, G., Geer, K., Krol, M., de Jong, P.  
and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished  
Other GSSs: RPCI-23-205J5-TV

## Contact: Shaying Zhao

## Department of Eukaryotic Genomics

## The Institute for Genomic Research

## 9712 Medical Center Dr., Rockville, MD 20850, USA

## Tel: 301 838 0200

## Fax: 301 838 0208

## Email: szhao@tigr.org

## Clones are derived from the mouse BAC library RPCI-23. For BAC

## library availability, please contact Pieter de Jong

## (pieter@dejong.med.buffalo.edu). Clones may be purchased from

## BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)

## or from Resea ch Genetics (info@resgen.com). BAC end page:

## http://www.tigr.org/tadb/bac\_ends/mouse/bac\_end\_intro.html

## Plate: 205 row: J column: 5

## Seq primer: T7

## Class: BAC ends.

## Location/Qualifiers

## 1. .454

## /organism="Mus musculus"

## /mol\_type="genomic DNA"

## /strain="C57BL/6J"

## /db\_xref="taxon:10090"

## /clone="RPCI-23-205J5"

## /sex="Female"

## /lab\_host="DH10B"

## /clone\_lib="RPCI-23"

## /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:

## EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or

## brain genomic DNA was isolated and partially digested

## with a combination of EcoRI and EcoRI Methyase. Size

## selected DNA was cloned into the pBACe3.6 vector at the

## EcoRI sites. The ligation products were transformed into

## DH10B electrocompetent cells (BRL Life Technologies)."

## BASE COUNT 86 a 134 c 95 g 139 t

## ORIGIN

## Query Match 66.7%; Score 20; DB 28; Length 454;

## Best Local Similarity 82.1%; Pred. No. 4.7e+02;

## Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

## QY 2 CGGATCCAGCGCCTCTGTTTGATGGC 29

## Db 172 CGGTTCCAGTGCCTCTGTTTGATGCC 199

## RESULT 13

## BF469784

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## COMMENT

BF469784 456 bp mRNA linear EST 04-DEC-2000  
UI-M-BH3-aut-d-03-0-UI.r1 NIH BMAP M.S4 Mus musculus cDNA clone  
UI-M-BH3-aut-d-03-0-UI 5', mRNA sequence.

AGENCOURT 456 bp mRNA linear EST 04-DEC-2000  
IMAGE:6399431 5', mRNA sequence.

EST. 456 bp mRNA linear EST 04-DEC-2000  
Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 456)  
Bonald, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery

Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548

Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA

Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mEST@mail.nih.gov

cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements

Seq primer: M13 Reverse.

# FEATURES

Location/Qualifiers

1..456  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-BH3-aut-d-03-0-UI"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NIH\_BMAP\_M\_S4"  
/note="Vector: pRT3D-Fac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH\_BMAP\_M\_S4 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged, normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH\_BMAP\_M\_S4, NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, NIH\_BMAP\_M\_S3.1, NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1. The subtracted library (NIH\_BMAP\_M\_S4) was constructed as follows: PCR amplified cDNA inserts from NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH\_BMAP\_M\_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"

BASE COUNT 82 a 132 c 95 g 146 t 1 others

Query Match 66.7%; Score 20; DB 10; Length 456;

Best Local Similarity 82.1%; Pred. No. 4.7e+02; Mismatches 5; Indels 0; Gaps 0;

QY 2 CCGGATCCAGCGCCTCTGTTTGATGCC 29

Db 212 CGGTTCCAGCGCCTCTGTTTGATGCC 239

# RESULT 14

D59983

LOCUS HUM079C11B Clontech human fetal brain polyA+ mRNA (#5335) Homo sapiens cDNA clone GEN-079C11 5', mRNA sequence.

ACCESSION D59983.1 GI:961622

VERSION D59983

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

# REFERENCE

1 (bases 1 to 215)

AUTHORS Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Takaichi, A., Takeda, S., Watanabe, T., Takahashi, E., Hirai, Y., Maekawa, H., Shin, S. and Nakamura, Y.

TITLE

Fujiwara et al. (1995)

# JOURNAL COMMENT

Unpublished  
Contact: Tsutomu Fujiwara  
Otsuka GEN Research Institute  
Otsuka Pharmaceutical Co., Ltd  
463-10 kagasuno Kawauchi-cho, Tokushima, Tokushima, 771-01 Japan  
Tel: 0886-65-2888  
Fax: 0886-37-1035.

# FEATURES

Location/Qualifiers

1..215  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GEN-079C11"  
/clone\_lib="Clontech human fetal brain polyA+ mRNA (#6535)"

BASE COUNT 51 a 54 c 54 g 56 t

# ORIGIN

Query Match 66.0%; Score 19.8; DB 14; Length 215;

Best Local Similarity 91.3%; Pred. No. 4.7e+02;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28

Db 27 ATCCAGCGCCTCTGTTTGATGG 49

# RESULT 15

Cl4671/c

LOCUS Cl4671

DEFINITION Cl4671 Clontech human aorta polyA+ mRNA (#6572) Homo sapiens cDNA clone GEN-079C11 3', mRNA sequence.

ACCESSION Cl4671

VERSION Cl4671.1 GI:1569378

KEYWORDS EST.

SOURCE Homo sapiens (human)

# ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 300)

AUTHORS Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Takaichi, A., Takeda, S., Watanabe, T., Takahashi, E., Hirai, Y., Maekawa, H., Shin, S. and Nakamura, Y.

TITLE Fujiwara et al. (1995)

# JOURNAL COMMENT

Unpublished  
Contact: Tsutomu Fujiwara  
Otsuka GEN Research Institute  
Otsuka Pharmaceutical Co., Ltd  
463-10 kagasuno Kawauchi-cho, Tokushima, Tokushima, 771-01 Japan  
Tel: 0886-65-2888  
Fax: 0886-37-1035.

# FEATURES

Location/Qualifiers

1..300  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GEN-079C11"  
/clone\_lib="Clontech human aorta polyA+ mRNA (#6572)"

BASE COUNT 70 a 80 c 80 g 67 t

# ORIGIN

Query Match 66.0%; Score 19.8; DB 14; Length 300;

Best Local Similarity 91.3%; Pred. No. 5.1e+02;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ATCCAGCGCCTCTGTTTGATGG 28

Db 189 ATCCAGCGCCTCTGTTTGATGG 167

Search completed: August 27, 2003, 09:30:45

Job time : 98.0378 secs



